Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
Department of Psychology, University of Michigan
Ann Arbor, Michigan

DAVID A. GRANT, Associate Editor
University of Wisconsin

WILLIAM K. ESTES, Associate Editor
Indiana University

Consulting Editors

NORMAN H. ANDERSON, *University of California, Los Angeles*
E. JAMES ARCHER, *University of Wisconsin*
FRED ATTNEAVE, III, *University of Oregon*
CLETUS J. BURKE, *Indiana University*
JAMES DEESE, *Johns Hopkins University*
PAUL M. FITTS, *University of Michigan*
FRANK A. GELDARD, *University of Virginia*
JAMES J. GIBSON, *Cornell University*
CLARENCE H. GRAHAM, *Columbia University*
HAROLD W. HAKE, *University of Illinois*
ARTHUR L. IRION, *Tulane University*
HOWARD H. KENDLER, *New York University*
DONALD B. LINDSLEY, *University of California, Los Angeles*
KENNETH MACCORQUODALE, *University of Minnesota*
QUINN MCNEMAR, *Stanford University*
LEO POSTMAN, *University of California, Berkeley*
L. STARLING REID, *University of Virginia*
KENNETH W. SPENCE, *State University of Iowa*
DELOS D. WICKENS, *Ohio State University*

CONTENTS

On Stimulus and Response Discriminability: H. P. BAHRICK AND M. NOBLE 449

An Investigation of Some Mathematical Models for Learning: C. F. FEY 455

The Empirical Determination of Game-Theoretical Strategies: H. KAUFMAN AND G. M. BECKER 462

Stimulus Generalization and Aggressive Verbal Stimuli: A. H. BUSS 469

Effect of Volume of Reinforcement and Number of Consummatory Responses on Licking and Running Behavior: H. L. SNYDER AND S. H. HULSE 474

Persistence of a Response in the Apparent Absence of Motivation: H. B. ROBINSON 480

Human Heart Rate Responses during Experimentally Induced Anxiety: G. E. DEANE .. 489

Effects of Density on Identification and Discrimination in Visual Symbol Perception: W. H. TEICHNER, R. REILLY, AND E. SADLER 494

Monetary Reward and Motivation in Discrimination Learning: L. B. MILLER AND B. W. ESTES 501

The Effect of Separation between Test and Inducing Fields on Brightness Constancy: B. DUNN AND H. LEIBOWITZ 505

Supplementary Report: Spatial Generalization of Voluntary Responses under Two Techniques of Study and Two Levels of Anxiety: B. D. DIXON AND D. D. WICKENS 508

Supplementary Report: Directional Effects and Sex in Kinesthetic Aftereffects: P. BAKAN, R. THOMPSON, AND G. WILDES 509

American Psychological Association

Vol. 61, No. 6 June 1961

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc., at Prince and Lemon Sts., Lancaster, Pa. The subscription rate per volume is $10.00, or $20.00 annually. Single copies are $2.00. Subscriptions, orders, address changes, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This Journal publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects, except when the latter are specifically oriented toward the extension of behavior theory. Experimental psychometric studies and studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology.

Normally, articles exceeding 20 printed pages cannot be accepted. Within this limit, the piecemeal experiment-by-experiment reporting of psychological research is discouraged, and the reporting of the data and conclusions of substantial segments of research efforts is encouraged. Specifically, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) *must* be presented in a single article. Special provision is made for the publication of Supplementary Reports and Replication Reports in articles of not more than 1.5 printed pages (see Editor's Note, this Journal, 1957, **53**, 1–2).

Address all articles submitted for publication to the Editor, Arthur W. Melton, Department of Psychology, University of Michigan, Ann Arbor, Michigan. **Manuscripts *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc., as described in the *Publication Manual of the American Psychological Association*.** When in doubt about practices of the Journal, authors should examine a recent issue.

Articles are published in the order of their receipt, except in rare circumstances. Authors are supplied with 50 free offprints without covers. All of the cost of an author's alterations in galley proof is charged to the author. Priority in publication is given to articles whose authors assume the full cost of publication. In 1961 the publication cost is $30 per printed page.

To save printing costs, supplementary material to articles in this Journal are deposited with the American Documentation Institute. Copies may be obtained by ordering the Document Number given in the footnote reference to ADI. Orders should be sent to ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. Orders *must* be accompanied by checks or money orders payable to the Chief, Photoduplication Service, Library of Congress.

ARTHUR C. HOFFMAN
Managing Editor

HELEN ORR
Promotion Manager

Second-class postage paid at Lancaster, Pa.

Send address changes to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must reach the Subscription Office by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

© 1961 by the American Psychological Association, Inc.

Journal of
Experimental Psychology

VOL. 61, NO. 6 JUNE 1961

ON STIMULUS AND RESPONSE DISCRIMINABILITY

HARRY P. BAHRICK AND MERRILL NOBLE
Ohio Wesleyan University *Kansas State University*

One way to view primary generalization is simply that it involves failure to discriminate. For our present purposes no theoretical account for this failure is needed. Following the traditional separation into stimulus and response generalization, the former would involve a failure to discriminate among stimuli (cf. Gibson, 1959) and may be observed by the customary procedure of employing a specific stimulus during acquisition and noting whether the response that occurred to that stimulus during acquisition also occurs when a different stimulus is presented on a test trial. Similarily, response generalization is viewed as involving failure to discriminate among responses. One way to test for response generalization is merely to note any occurrence of response variation during the acquisition of a given specific response (Brogden, 1951, p. 590). Moreover, if the various responses so observed are continuously variable along a single dimension, the obtained distributions will reveal the statistical nature of response discriminability.

If instead of presenting a single S-R situation, *S* is required to produce to different stimuli responses that differ with regard to their position along some continuum, overlap among the response distributions may occur. The amount of this overlap would be dependent upon the statistical nature of the response distributions and the distance between adjacent required response points on the continuum. Noble and Bahrick (1956) studied response distributions in such a situation and found that closer spacing of required response points resulted in less response variability than did the wider spacing of required response points. If the response distributions are viewed as empirical gradients of response generalization, this result may be restated as indicating that amount of response generalization is inversely related to increase in response similarity. Another result of the earlier study, however, was that despite this effect adjacent response distributions overlapped more with the more similar responses than with the wider response spacing. Since the more similar responses were selected from a previously determined psychophysical scale to be sufficiently close that some overlap would be expected, the latter result presumably reflects some fundamental limi-

tation in *S*'s ability to discriminate.

In the earlier investigation (Noble & Bahrick, 1956) there was no special attempt to scale the stimuli of the task. In the present study both stimuli and responses were scaled in an attempt to determine the relative effects of stimulus and response similarity upon response distributions. The investigation was carried out in two parts. In Part I a psychophysical scale of stimulus similarity was developed. From this scale two sets of stimuli, differing in amount of separation from one another, were employed in a factorial design in combination with two sets of responses determined by the psychophysical scale of the earlier study.

Part I

Method

Apparatus.—The apparatus was essentially a device to expose a horizontal black line, ¼ in. wide and always centered on a white background. The *S* viewed this line through an aperture 1 in. × 8 in., and either *S* or *E* could vary the length of the black line from 0 in. to 5 in. by means of a control knob located on either side of the apparatus. The device was equipped with a shutter to permit exposure of the black line for varying intervals.

Subjects.—Ten male undergraduates at Ohio Wesleyan University were employed on an hourly basis. Each *S* served under each of 10 stimulus conditions.

Stimuli.—The following 10 lengths of lines, in inches, were used as stimuli: .250, .375, .500, .750, 1.000, 1.250, 1.500, 1.750, 2.000, and 2.500.

Procedure.—Each *S* made a total of 80 responses to a single one of the 10 stimuli during each of 10 half-hour sessions. The 10 sessions were separated from one another by at least 24 hr. The sequence of stimuli presented to each *S* in successive sessions was counterbalanced by means of a random latin square. The *S* was seated so that the exposure device was at eye level and at a distance of 24 in. The particular length of line previously selected was then exposed for a period of 4 sec. At the end of this interval the shutter closed and the line was reduced to zero length. The shutter was then opened again and *S* was given 4 sec. to reproduce the same length of line as previously exposed as accurately as possible by manipulating his control knob. All responses were recorded by *E* by means of a scale which could be read to the nearest .001 in. Trials were given at 15-sec. intervals.

Results

Analyses of the data showed no appreciable learning beyond the first few responses to any stimulus. Accordingly, means and *SD*s of the distributions of responses were determined for each *S* for each stimulus for Trials 11–80. Constant errors were small and normally distributed. Average *SD*s were computed for each stimulus length, and a scale of equal discriminability was obtained by a procedure described earlier (Noble & Bahrick, 1956). From this scale successive stimulus points were selected so that they were separated by the square root of the sum of squares of their respective average *SD*s. The shortest length was arbitrarily chosen at .250 in.; the successive stimuli are as follows (all values are in inches): 0.250, 0.347, 0.451, 0.559, 0.673, 0.794, 0.921, 1.056, 1.198, 1.348, 1.506, 1.673, and 1.850. The resulting scale approximates an equal interval scale, since there should be an approximately equal percentage of overlap of response distributions for all adjacent points on the scale.

Part II

Method

Apparatus.—The apparatus consisted of a memory drum for the presentation of stimuli, and of a semirigid pressure control (Noble & Bahrick, 1956) for the responses.

Subjects.—Forty-eight right-handed male students who had not participated as *S*s in Part I were divided by means of a table of random numbers into four groups of 12 *S*s each.

Procedure.—The *S*s were seated at the pressure control and a memory drum was

mounted so that its 1-in. × 8-in window was located at eye level and at a distance of 24 in. from *S*. The drum exposed, at intervals of 8 sec., horizontal black lines $\frac{1}{16}$ in. wide, centered against a white background, and of a length corresponding to those shown in Table 1 for each group of *S*s. The sequence of the seven alternative stimuli for each group was drawn on a tape in an order determined by a table of random numbers. In any sequence the order did not repeat itself for 70 stimuli, and was chosen with the constraint that each stimulus occurred 10 times.

Subjects were instructed to respond to the stimuli by exerting force on the control; the amount of force to be an increasing function of the length of the line exposed. The *S*'s responses were recorded to the nearest $\frac{1}{4}$ lb., and knowledge of results was given by informing *S* how many $\frac{1}{4}$-lb. units he had overshot or undershot the correct force assigned to that particular stimulus. Thus, if *S* responded with a pressure of 45 quarter-pounds to a line of 0.921 in., *E* said "6 over," since the correct response was 39 quarter pounds for all groups. Or, if an *S* of Group WSWR responded with a force of 14 quarter-pounds to a line of 0.451 in., *E* said "2 over," since the correct response was 12 quarter-pounds for *S*s of this group.

At the end of the 8-sec. exposure, the drum exposed a new stimulus and *S* continued as before, to give a total of 280 responses (40 to each of the seven stimuli assigned to his group).

S-R conditions.—The stimulus and response conditions for the four groups are summarized in Table 1. The successive points on the equal discriminability stimulus scale are from Part I; the equal discriminability scale of responses is from our earlier study (Noble & Bahrick, 1956), rounded to the nearest $\frac{1}{4}$ lb. Group WSWR was trained with wide spacing of stimuli and wide spacing of responses by selecting alternate points on the stimulus and response scales. Group WSNR was trained with wide spacing of stimuli, but narrow spacing of responses. This was done by choosing the same alternate points on the stimulus scale, but the seven central and adjacent points on the response scale. Group NSWR was trained with narrow stimulus spacing and wide response spacing by selecting the seven central stimulus points, but alternate response points, and Group NSNR was trained with the seven central points on both scales.

Results

Mean constant errors and the square root of the within-group *MS* were computed, for each group separately, for successive five-trial blocks for each type of response. As in the previous study (Noble & Bahrick, 1956) little learning occurred beyond the first 20 responses to any stimulus, and therefore all subsequent comparisons are based upon Responses 21–40 to each stimulus.

The constant errors are shown in Fig. 1. It can be seen that all four groups in the present study showed a definite range effect (Ellson & Wheeler, 1949). The *S*s tended to respond with too much force at the lower end (positive constant errors) and with too little force (negative constant error) at the high end of the force continuum. This tendency was present to a lesser extent in the earlier study (Noble & Bahrick, 1956). In that study the two groups using

TABLE 1

STIMULUS AND RESPONSE VALUES FOR THE FOUR CONDITIONS

Cond.	Stimulus Scale (in inches)												
	.250	.347	.451	.559	.673	.794	.921	1.056	1.198	1.348	1.506	1.673	1.850
WSWR	4		12		24		39		58		85		119
WSNR	18		24		31		39		48		58		71
NSWR				4	12	24	39	58	85	119			
NSNR				18	24	31	39	48	58				

Note.—Values in the cells represent required response force, in $\frac{1}{4}$-lb. units.

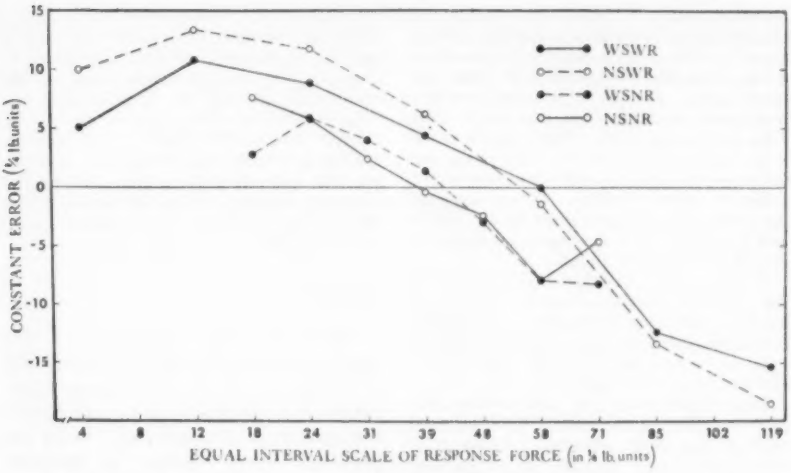

FIG. 1. Mean constant error. (The mean differences between correct force and force actually exerted are shown for the last 20 responses to each stimulus.)

highly discriminable stimuli generally showed positive constant errors at the end of training except for the very highest response values. Inspection of Fig. 1 also shows that the groups with wide response spacing showed larger constant errors than the groups with narrow spacing between responses. Analysis of variance shows, however, that the differences among the groups just failed to reach significance ($P > .05$) at the three common response points of 24, 39, and 58 quarter-pounds.

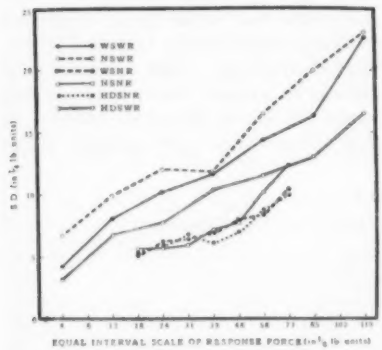

FIG. 2. Variable error. (Results Cond. HDSNR and HDSWR of a previous study —Noble & Bahrick, 1956—with highly discriminable stimuli and the same two sets of responses as the present study are included.)

Figure 2 shows the square root of the mean within-group variance for the four groups of the present study. Also shown are results from the previous study for the two groups using highly discriminable stimuli and the same two sets of responses as in the present study. These latter two groups are identified as HDSNR and HDSWR, respectively.

Two conclusions are at once apparent from inspection of Fig. 2. First, all groups with wide response spacing have larger response variance than the groups with narrow response spacing, and second, stimulus spacing does not have much effect for the groups with narrow response spacing, but does have an effect for the groups with wide response spacing. Comparisons by means of t tests of the

mean within-groups response variances between the two groups with the more widely spaced responses (Groups WSWR and NSWR) show significant ($P < .05$) differences for four out of seven comparisons and all seven differences are in the same direction. A similar comparison between the two groups with narrow response spacing (WSNR and NSNR) shows that none of the seven differences are significant ($P > .05$), and there are two reversals.

The average amount of response information (I_R), response equivocation (E_R), and information transmitted (I_T) was computed for the four groups by the same method as in the earlier study (Noble & Bahrick, 1956). The results of these analyses, together with those of the earlier study, are shown in Table 2. These results also suggest that overlapping among response distributions was not materially increased by narrower stimulus spacing if the response spacing was also narrow. The amount of response equivocation, which reflects this overlapping, varies only by a small amount for the two groups involved. For the groups with wide response spacing, however, this is not true. Here the increased stimulus similarity resulted in more response equivocation, so that the transmitted information was reduced.

TABLE 2

INFORMATIONAL ANALYSES FOR THE LAST 20 RESPONSES TO EACH STIMULUS, SHOWING, IN BITS, AVERAGE AMOUNT OF RESPONSE INFORMATION (I_R), RESPONSE EQUIVOCATION (E_R), AND INFORMATION TRANSMITTED (I_T)

Condition	I_R	E_R	I_T
HDSWR	6.775	5.188	1.587
WSWR	6.700	5.360	1.340
NSWR	6.714	5.410	1.304
HDSNR	6.010	4.920	1.090
WSNR	6.003	4.933	1.070
NSNR	5.988	4.951	1.037

DISCUSSION

It is apparent that response similarity has a greater effect upon response generalization under the conditions of our investigation than does stimulus similarity. This comparison is possible since the scales of similarity for stimuli and responses were developed by comparable operations. Wider spacing among responses leads to a greater amount of absolute generalization, i.e., to less steep generalization gradients, but also to a smaller amount of relative generalization; that is, to less overlap among adjacent response distributions. Most studies of response generalization obtain, in effect, only the latter type of measure. The usual measure of generalization in studies employing discrete responses is the number of intrusions, and these reflect response overlap which is a joint function of the steepness of the generalization gradients and of the interval distance between response points. Only where responses can be scaled along a continuum is a separation of these factors possible.

The present study suggests that stimulus spacing has only a minor effect upon response generalization gradients when the stimuli are no less discriminable than the responses, and if the intervals between required responses are small. This is particularly apparent if one takes into account the record of the VWSNR group from the previous study (Noble & Bahrick, 1956) as well as the present results. Overlap among response distributions was not significantly affected by variations of stimulus similarity from the highly discriminable to the barely discriminable.

When responses are more highly discriminable, on the other hand, stimulus similarity begins to have a more pronounced effect upon response distributions. With more discriminable stimuli, constant errors, particularly those that reflect the range effect, are minimized and generalization gradients become some-

what steeper so that overlap among response distributions is somewhat reduced. The effect upon the gradients remains smaller, however, than that produced by comparable response spacing. These conclusions are limited to the unidimensional variation of responses which was employed in the present study. Such variation limits the degree of response discriminability, and it remains to be seen whether multidimensional variation of responses will produce similar findings.

Summary

Subjects were trained to exert scaled amounts of force on a semirigid control stick in response to scaled lengths of lines presented on a memory drum. Four task versions were used, representing two degrees each of stimulus and response discriminability.

The response distributions obtained to each stimulus are interpreted as gradients of primary response generalization. It was shown that the steepness of these gradients was greater with narrow spacing among the required responses. Thus greater similarity of responses diminished the absolute generalization, but nevertheless increased the overlap among adjacent response distributions. Stimulus spacing had comparatively little effect upon response generalization. Great variation of stimulus spacing produced no significant changes upon response generalization when the required responses were spaced closely together, and when the stimuli were no less discriminable than the responses.

Emphasis was placed upon the distinction between absolute amounts of generalization, as reflected by direct examination of the generalization gradients, and measures of relative generalization such as "intrusions." Such measures reflect overlap among response distributions and are a function of both absolute generalization and the spacing between the required responses.

References

Brogden, W. J. Animal studies of learning. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 568–612.

Ellson, D. G., & Wheeler, L., Jr. The range effect. *USAF Air Materiel Command tech. Rep.*, 1949, No. 5813.

Gibson, E. J. A re-examination of generalization. *Psychol. Rev.*, 1959, **66**, 340–342.

Noble, M. E., & Bahrick, H. P. Response generalization as a function of intratask response similarity. *J. exp. Psychol.*, 1956, **51**, 405–412.

(Received April 25, 1960)

Journal of Experimental Psychology
1961, Vol. 61, No. 6, 455-461

AN INVESTIGATION OF SOME MATHEMATICAL MODELS FOR LEARNING[1]

CURT F. FEY[2]

University of Pennsylvania

In this study an attempt is made to determine whether the results of two different learning experiments can be described by stochastic models proposed by Bush and Mosteller (1955) and Luce (1959) without changing the model parameters.

The merit of a model lies in its ability to describe and predict data successfully with the aid of a minimum of free parameters. For any one experiment this can be done with several models. Consequently a more stringent test of a model is its ability to predict the fine structure of the data with one invariant set of parameters in such a way that once the values of the parameters are determined in one experiment these same parameters can be used to predict the outcome of another experiment.

Galanter and Bush (1959) previously studied parameter invariance in the linear model of Bush and Mosteller (1955). Their analysis showed an apparent lack of parameter invariance in a T-maze situation, but it is not clear whether the lack of parameter invariance was attributable to a basic mechanism in the model or was a consequence of sampling errors and difficulties in estimating parameters.

The purpose of the present study is to investigate this question of parameter invariance in greater detail. The experimental design was improved over that of Galanter and Bush (1959) by running only one rather than three trials per day, and it was extended to provide a comparison between 100% reinforcement and 75% random reinforcement.

The Models

The two models used in this paper may be designated as the alpha model (Bush & Mosteller, 1955) and the beta model (Luce, 1959). Each of them uses linear transformations. In the alpha model the linear transformation is applied to the response probability p; in the beta model it is applied to the quantity $p/(1-p)$.

Both of these models are stochastic, i.e., they deal with probabilities of making responses. The models are path-independent: the response probability on a given trial depends only on the response probability and the outcome on the previous trial.

An animal in a T maze can turn either to the left or to the right on any given trial. The models state that if on one trial S makes a response for which it gets rewarded, then the probability of making that same response on the next trial increases. The models specify the manner of these changes.

Let p_n be the probability of going to the right-hand side (probability of an "error") on trial n; let $q_n = 1 - p_n$; let α_1, α_2, β_1 and β_2 be nonnegative parameters such that α_1 and β_1 are associated with reward and α_2 and β_2 are associated with nonreward. The models can then be defined in the following way:

[1] Based on the author's PhD dissertation supervised by R. R. Bush, and read by R. D. Luce and J. Beck. The data analysis was performed at the Computer Center of the University of Pennsylvania with the assistance of S. Gorn and P. Z. Ingerman.

[2] Now at General Dynamics/Electronics, Rochester, New York.

Alpha Model	*Response*	*Outcome*	*Beta Model*
$p_{n+1} = \alpha_1 p_n$	left turn	reward	$p_{n+1} = \frac{p_n}{p_n + \beta_1(1 - p_n)}$
$p_{n+1} = \alpha_2 p_n$	right turn	nonreward	$p_{n+1} = \frac{p_n}{p_n + \beta_2(1 - p_n)}$
$q_{n+1} = \alpha_1 q_n$	right turn	reward	$p_{n+1} = \frac{\beta_1 p_n}{\beta_1 p_n + 1 - p_n}$
$q_{n+1} = \alpha_2 q_n$	left turn	nonreward	$p_{n+1} = \frac{\beta_2 p_n}{\beta_2 p_n + 1 - p_n}$

Mathematical properties of the alpha model were listed by Galanter and Bush (1959, pp. 272–273) for the special condition that a left response is always rewarded and a right-hand response is never rewarded, (100:0). For the beta model the mathematical properties have been determined by Kanal (1960, 1961), Bush, Galanter, and Luce (1959, p. 387), and Bush (1960).

METHOD[1]

. *Subjects.*—The *S*s were male hooded rats of the Long-Evans strain, from Rockland Farms, New York City, New York. They weighed about 75 gm. on arrival. Eight rats were used for the preliminary experiment. In the main experiment 63 rats were used, but the final $N = 50$, because 13 died during the experiment.

[1] For details see Fey, 1960.

TABLE 1

PERIOD 2, 100:0 GROUP

COMPARISON OF STATISTICS FROM THE FIRST 35 TRIALS OF EXPERIMENTAL GROUP WITH CORRESPONDING MODEL VALUES CALCULATED WITH $p_1 = 1$, $\alpha_1 = .858$, AND $\alpha_2 = .955$ FOR ALPHA MODEL AND $p_1 = .97$, $\beta_1 = .952$, AND $\beta_2 = .642$, FOR BETA MODEL

Statistic	Means			Standard Errors		
	Exp.	α Model	β Model	Exp.	α Model	β Model
Number of *S*s	25	100	500			
Number of trials	35	35	35			
Total number of errors	12.28	12.28	12.39	.76	.00	.012
Trial of last error	23.16	22.10	26.45	1.00	.25	.028
Trial of first success	6.88	6.87	6.49	.48	.00	.024
Number of RR sequences	7.48	7.32	7.01	.56	.00	.014
RL	4.76	4.85	5.32	.28	.09	.016
LR	3.80	3.85	4.41	.28	.10	.014
LL	17.96	17.98	17.26	.80	.22	.024
Number of L runs of:						
Length 1	2.00	1.81	1.86	.20	.10	.012
2	.56	.81	.87	.08	.03	.014
3	.44	.51	.57	.08	.03	.006
4	.32	.37	.38	.04	.04	.006
5	.08	.28	.30	.04	.02	.004
Number of R runs of:						
Length 1	2.60	2.70	3.12	.24	.00	.014
2	.80	.81	.83	.16	.00	.008
3	.40	.40	.36	.08	.00	.006
4	.24	.27	.28	.04	.00	.008
5	.20	.20	.20	.04	.00	.004
Total number of R runs	4.80	4.96	5.38	.28	.00	.018

Note.—Standard error of the mean was computed from range approximation.

TABLE 2

PERIOD 2, 75:25 GROUP

COMPARISON OF STATISTICS OF 75:25 EXPERIMENTAL GROUP WITH CORRESPONDING MODEL VALUES CALCULATED WITH $p_1 = 1$, $\alpha_1 = .858$, AND $\alpha_2 = .955$ FOR ALPHA MODEL AND $p_1 = .97$, $\beta_1 = .952$, AND $\beta_2 = .642$ FOR BETA MODEL

Statistic	Means			Standard Errors		
	Exp.	α Model	β Model	Exp.	α Model	β Model
Number of animals	25	100	200			
Total number of errors	15.92	19.21	19.43	1.04	.22	.03
Trial of last error	27.32	32.71	33.59	1.04	.21	.045
Trial of first success	7.52	7.19	7.56	.64	.16	.07
Number of RR sequences	10.24	11.85	11.81	.88	.25	.045
RL	5.44	7.09	7.33	.28	.10	.035
LR	4.68	6.37	6.75	.32	.09	.035
LL	13.64	8.69	8.11	1.12	.24	.08
Number of L runs of:						
Length 1	2.48	3.63	3.55	.20	.10	.045
2	1.16	1.62	1.68	.20	.05	.03
3	.40	.77	1.09	.08	.03	.02
4	.24	.41	.52	.04	.02	.015
5	.16	.23	.21	.04	.02	.01
Number of R runs of:						
Length 1	2.88	3.52	3.70	.20	.10	.035
2	1.16	1.50	1.63	.12	.05	.03
3	.48	.66	.79	.12	.04	.015
4	.36	.54	.47	.08	.03	.02
5	.20	.38	.29	.08	.02	.01
Total number of R runs	5.68	7.36	7.62	.32	.09	.035

Note.—The model parameters were estimated from the 100.0 group. Standard error was computed from range approximation.

Apparatus.—The T maze was a replica of that used by Galanter and Bush (1959). It consisted of a straight alley runway for pretraining and a T maze for the main experiment. The T maze was built in such a way that the crossbar and the start arm of the T could be separated and a goalbox could be hooked to the stem of the T, thereby changing the maze into a straight runway. The maze was built of plywood with a removable wire mesh top and pressed wood doors. The inside of the stem and the attachable goalbox were painted medium gray, the right arm was painted light gray, and the left, dark gray. The length of the cross arm was 60 in., the length of the stem was 26 in., and the attachable goalbox was 10 in. The alleys were 4 in. wide and the walls were 8 in. high. The starting compartment was 10 in. long with a guillotine door on the maze side and a hinged door on the outside. Another guillotine door was at the choice point. The goal cups were placed at the end of each arm. The metal goal cups had double floors, the bottom part contained inaccessible wet food mash to balance olfactory cues, and the top contained the reward pellet.

Procedure.—This experiment consisted of three parts: (*a*) preliminary handling; (*b*) straight alley pretraining; and (*c*) T-maze learning.

The *S*s were kept in the laboratory for 23 days at ad lib. food and water and were handled daily. Then *S*s were deprived of food for 18, 21, 21¼, 21¾, and 22 hr. on Days 24, 25, 26, 27, and 28, respectively.

The pretraining started on Day 29. For the remainder of the experiment *S*s were under 18 hr. food deprivation at the beginning of each daily run. They were fed 4 hr. later for a 2-hr. period. Water was always available in the cages.

The *S*s were given one trial per day of pretraining on the straight alley runway. Pretraining lasted for three days.

During the 30 days of Period 1 of the T-maze learning, the following procedure was adhered to: .038-gm. pellet was deposited in the right goal cup; nothing was placed in the left goal cup. The *S* was placed in the

startbox and the startbox door was raised. As *S* passed the choice point, its door was lowered. The *S* was left in the maze until it ate the pellet, until it investigated the goal cup (on the nonrewarded side), or until 3 min. were up, whichever occurred first.

At the end of Period 1 *S*s were divided at random into two groups. One group was always rewarded on the left side during Period 2, and the other was rewarded accoring to the following schedule obtained from a random number table with $P(L) = 0.75$: L L L R L L R L L R L L L L L L R L R L L L L R L L R R R R L L L L L. Period 2 lasted for 35 days.

Estimation of parameters.—The parameters of the alpha model were estimated in the following way: The initial probability p_1 was taken to be 1.00. The other two parameters were estimated from the Period 2 data of the 100:0 group by equating the observed mean number of trials before the first success and the observed mean total number of errors to their respective expected values.

Initial estimates of the beta model parameters were determined by methods similar to those used for finding the alpha model parameters. These estimates were modified by exploration of the parameter space until the response probabilities (Monte Carlo computations) were similar to the experimental data of the 100:0 group. The following criteria were used: the total number of errors generated by the model had to match the data, and a plot of trial-by-trial mean response probabilities produced by the model had to appear similar to the corresponding plot of the data.

Results

The results of the experiment are summarized in Fig. 1 and 2 and Tables 1 and 2. Figure 1 presents the proportions of R response of the 100:0 group during Period 2 and the corresponding curves generated by the models. Figure 2 depicts the same data for the 75:25 group during Period 2. Tables 2 and 1 give

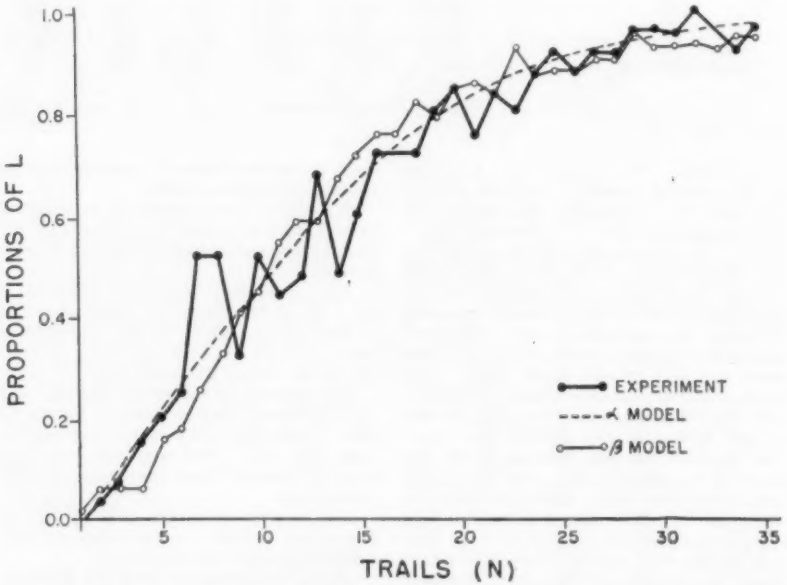

FIG. 1. Period 2, Group 100:0. Trial by trial proportions of L responses made by 25 experimental *S*s (filled circles); generated by alpha model (smooth line) computed with $p_1 = 1.00$, $\alpha_1 = 0.858$, and $\alpha_2 = 0.955$; and generated by 500 beta model Monte Carlo analogs (open circles) computed with $p_1 = 0.97$, $\beta_1 = 0.952$, and $\beta_2 = 0.647$.

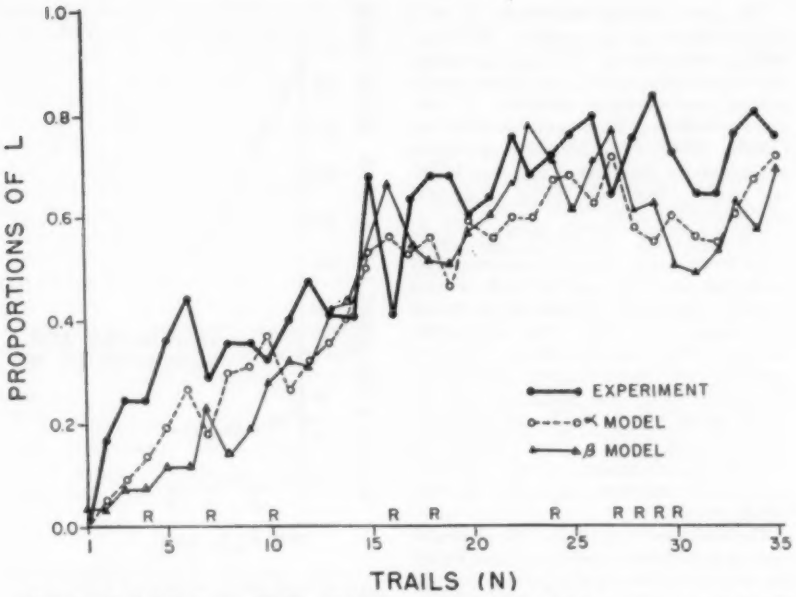

FIG. 2. Period 2, Group 75:75. Trial by trial proportions of L responses made by 25 experimental *Ss* (filled circles); by 100 alpha model Monte Carlo analogs (open circles) computed with $p_1 = 1.00$, $\alpha_1 = 0.858$, and $\alpha_2 = 0.955$; and by 200 beta model Monte Carlo analogs (triangles) computed with $p_1 = 0.97$, $\beta_1 = 0.952$, and $\beta_2 = 0.647$. (R = food reward is in right maze arm, otherwise the left arm is baited.)

comparative results of this experiment and corresponding model values. A more detailed analysis of results is presented by Fey (1960).

DISCUSSION

The merit of a mathematical model of learning lies not so much in describing the data of any one experiment with the aid of parameters estimated from that particular experiment as in its ability to represent accurately the learning process of a variety of different experimental situations using the *same* set of parameters. In other words once the parameters are estimated for one experimental situation the model should be able to predict the course of learning in other experiments. Models which will handle a variety of experimental situations with the same set of parameters are called *parameter invariant*.

This experiment indicates that the models under consideration fit the Period 2, 100:0 group data, from which their parameters were estimated, quite well, but the fit to the Period 2, 75:25 group data (using parameters estimated from the Period 2, 100:0 group) is less successful. Both models show an apparent lack of parameter invariance of approximately equal magnitude.

Tables 1 and 2 might give the impression that the alpha model fits the data slightly better than the beta model. This conclusion is hardly warranted if the magnitudes of the differences and the methods of estimating the parameters are considered. The alpha model parameters were determined analytically; those of the beta model were estimated by Monte Carlo procedures. Thus the alpha model parameters were determined more exactly than those of the beta model.

The lack of long runs seems to be a basic difficulty of the models. This is of little consequence in 100:0 animal learning, but it does seem to be important in partial reinforcement schedules for animals as well as in human choice behavior (Derks, 1960). This lack of long runs is not generally manifested in mean learning curves, but only in a sequential analysis of the data.

The fact that the 75:25 "stat rats" learn more slowly than the experimental *Ss* is not as serious as the lack of long runs. A change in the size of the model parameters will correct the former deficiency. The data indicate that by reducing the beta model parameters by about 25%, the total number of errors made by the model analogs will match those of the experimental *Ss* for the 75:25 group. These reduced parameters decrease the fit to the 100:0 group.

The slow learning of the 75:25 model analogs could be handled by specifying the manner in which the parameters are modified when the schedule changes from 100:0 to 75:25. With respect to the lack of perseverance, no small change in parameter values would increase the fit of model to data.

Galanter and Bush (1959) noted in three of their experiments that the probability of turning to the more frequently rewarded side tended to decrease slightly during the first few acquisition trials before it began to rise. This phenomenon occurs also in other experimental situations (Gibson & Walk, 1956; Jensen, 1960; Kendler & Lachman, 1958). In the present experiment, the initial dip is hardly noticeable.

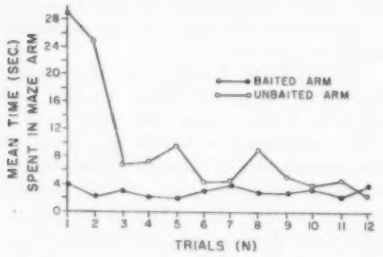

FIG. 3. Trial by trial distribution of time spent in baited, left (filled circles) and unbaited, right (open circles) maze arm by 20 *Ss* of Galanter and Bush (1959) Exp. III.

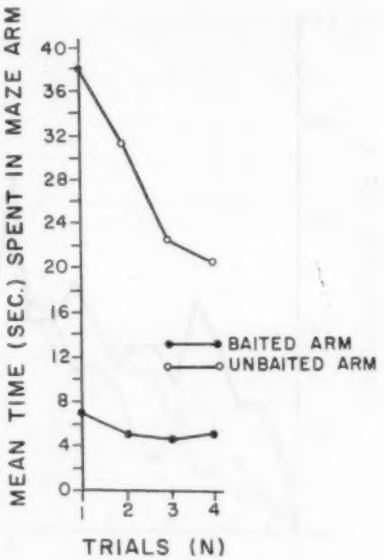

FIG. 4. Trial by trial distribution of time spent in baited, left (filled circles) and unbaited, right (open circles) maze arm by 50 *Ss* of Period 1.

A look at the time the rats spent in the baited and in the unbaited arms of the maze (Fig. 3 and 4) indicates that initially our *Ss* and those of the Galanter and Bush (1959) Exp. III[4] were removed more quickly from the unbaited than from the baited side of the maze; later in the experiment, removal occurred after approximately the same time interval in either arm of the maze. The reason for this is found in the criteria for removing *S* from the maze: *S* is left in the maze until it investigates the food cup on the unbaited side, until it eats the pellet on the baited side, or until 3 min. are up, whichever occurs first. The *Ss* investigate the food cup on the

[4] The data plotted in Fig. 3 were obtained from the original protocols of the experiment reported by Galanter and Bush (1959).

baited side before they start eating the pellet. In fact, *S* may take a pellet in its mouth, drop it, and not eat it; thus investigation of the food cup occurs before the eating (Fig. 3 and 4).

Should *S* initially prefer removal from the maze to eating the pellet, the nonrewarded side may actually be more attractive to *S* than the rewarded side, since *S* has to stay in the nonrewarded side for a shorter period of time. The *S*s behave initially as if they were much more interested in escaping from the maze than in eating. As *S* becomes accustomed to the experimental situation, interest in escaping decreases and the food pellet gradually becomes more attractive.

This explanation can handle the dip in our experiment, but it fails in the case of other experiments such as a Skinner-box situation.

Summary

This paper investigated two models for learning: a linear model proposed by Bush and Mosteller and a nonlinear model developed by Luce. Specifically, an attempt was made to determine whether data obtained from two different experimental situations could be described by the models without changing the parameters.

Two groups of rats were trained in a T maze. One group was always rewarded with food on one side; the other group received a food reward with probability .75 on one side and .25 on the other side. Model statistics were computed for both groups, using parameters estimated from the group that was always rewarded on the same side, and compared with the experimental data.

It was found that there is good agreement between the models and the data of the continuously reinforced group, from which the model parameters were estimated. The fit to the data of the partially reinforced group, however, leaves something to be desired.

Both models fit the data about equally well.

References

Bush, R. R. Some properties of Luce's beta model for learning. In K. J. Arrow, S. Karlin, & P. Suppes (Eds.), *Proceedings of the first Stanford symposium on mathematical methods in the social sciences.* Stanford, Calif.: Stanford Univer. Press, 1960.

Bush, R. R., Galanter, E., & Luce, R. D. Tests of the "beta model." In R. R. Bush & W. K. Estes (Eds.), *Studies in mathematical learning theory.* Stanford, Calif.: Stanford Univer. Press, 1959.

Bush, R. R., & Mosteller, R. *Stochastic models for learning.* New York: Wiley, 1955.

Derks, P. Human binary prediction and the "conditioning axiom" under temporal, incentive, contingency, and experimental variations. Unpublished doctoral dissertation, University of Pennsylvania, 1960.

Fey, C. F. Parameter invariance in models for learning. Unpublished doctoral dissertation, University of Pennsylvania, 1960.

Galanter, E., & Bush, R. R. Some T-maze experiments. In R. R. Bush & W. K. Estes (Eds.), *Studies in mathematical learning theory.* Stanford, Calif.: Stanford Univer. Press, 1959.

Gibson, E. J., & Walk, R. D. The effect of prolonged exposure to visually presented patterns on learning to discriminate them. *J. comp. physiol. Psychol.*, 1956, **49**, 239–242.

Jensen, G. D. Learning and performance as functions of ration size, hours of privation, and effort requirement. *J. exp. Psychol.*, 1960, **59**, 261–268.

Kanal, L. Analysis of some stochastic processes arising from a learning model. Unpublished docotral dissertation, University of Pennsylvania, 1960.

Kanal, L. On a random walk related to a nonlinear learning model. *IRE Nat. convention Rec.*, 1961, in press.

Kendler, H. H., & Lachman, R. Habit reversal as a function of schedule of reinforcement and drive strength. *J. exp. Psychol.*, 1958, **55**, 584–591.

Luce, R. D. *Individual choice behavior: A theoretical analysis.* New York: Wiley, 1959.

(Early publication received December 6, 1960)

Journal of Experimental Psychology
1961, Vol. 61, No. 6, 462–468

THE EMPIRICAL DETERMINATION OF GAME-THEORETICAL STRATEGIES[1]

HERBERT KAUFMAN AND GORDON M. BECKER[2]

Electric Boat Division, General Dynamics Corporation

The present paper reports a study of the behavior of naive subjects in a strategy-making situation where a normative solution, i.e., game theory, can be stated. In simplest game theory terms, *Ss* were required to find optimal strategies in a two person zero sum game under conditions to be described. The game theory solution provides a logical reference or ideal to which the actual game playing is compared. The procedure provides a dynamic behavioral analysis of decision problems usually handled by the game theory model.

Atkinson and Suppes (1958, 1959), Lieberman (1959), and Simon (1956) have previously reported empirical studies involving game theory concepts. These studies, almost universally, are concerned with behavior in game situations where individual moves, rather than strategies, are made at each trial. The present paper represents a radical departure from this usual procedure in that the individual playing a matrix game is required to indicate a game strategy at every trial, i.e., to select 100 moves at each trial, rather than simply one move.

In order to explain the experimental procedure, some characteristics of the game theory model will be stated: (*a*) In the two person, two choice, zero sum game represented by a 2×2 matrix there exist optimal strategies for each player. (*b*) If either player plays his optimal strategy the play of the other has no effect on the average payoff. (*c*) Playing an iterated game involves, for each player, both a decision about the relative frequencies of play for the two choices, the game *strategy*, and a decision about the sequence of play, the game *tactics*. In the absence of any predictive criteria on the play of the opponent, a random sequence of the choices is to be preferred in order to guarantee the expected payoff for a given strategy.

In the experiment to be described, the strategy-making aspect of the game decisions was isolated by removing the burden of tactical play from *Ss*. The major questions to be answered were: (*a*) Do naive *Ss* learn, through the experience of playing a game, to find and maintain a strategy, optimal or otherwise? (*b*) Do games having the same values (average payoff for optimal strategy), but different optimal strategies, differ in ease of learning?

METHOD

Strategy decisions were emphasized by forcing *Ss* to choose relative frequencies of play, and not allowing them to decide upon the sequence of play. At each trial *S* divided 100 choices between his two alternatives. Each block of 100 choices amounts to a strategy. Thus, at each trial *S* announced the strategy he wanted on that trial, e.g., 42 of one alternative and 58 of the other. The *Ss* strategy was paired with a strategy chosen by *E* to be punishing to *S* (see below). That is to say, in every game *S*'s opponent

[1] This research was done as part of the Basic Research Program of General Dynamics Corporation. The authors wish to thank Mortimer H. Applezweig for his help in providing Connecticut College psychology laboratory facilities and students.

[2] Now at the University of California, Los Angeles.

was in fact *E*. The payoff to *S* for this pair of opposing strategies was calculated under an assumption of random play by both players, equivalent to arranging the two sequences of choices each in random order and summing the 100 payoffs thus determined.

Figure 1 is the graphical representation of Game D shown in Table 1. T and B are the alternative choices for the row player, L and R the alternatives for the column player, and the cell entries represent the payoffs to the row player for the given combinations of row and column choices. In this figure the strategies of the row player are shown along the abscissa, labeled as the proportion of top row (T choices). For example, the point 92 on the abscissa corresponds to the row player's choice of 92T and 8B. The ordinate is payoff points to the row player. The four ordinate values, *a*, *b*, *c*, and *d*, are the cell entries of the matrix in the top row (*a*, *b*) and the bottom row (*c*, *d*), from left to right. For example, in Game D (Table 1), *a* is 1.12, *b* is 0.78, *c* is 0.45, and *d* is 1.47. The column strategies are represented parametrically by the lines cutting across from the line X = 0 to the line X = 100, labeled as left column proportion. For example, the line running from (100, 85) to (0, 126) corresponds to the column player's strategy of 20L and 80R. The payoff to the row player for 92T–8B against a column strategy of 20L–80R is found on the graph to be 88 points.

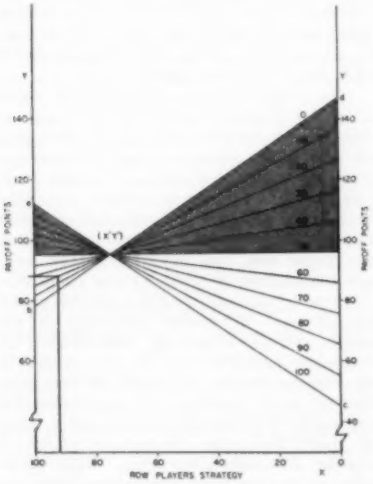

FIG. 1. Graphical illustration of Game D.

TABLE 1

FIVE EXPERIMENTAL GAMES

Game	*S* Chooses Rows T & B	Payoffs to *S* When *E* Chooses Column		Optimal Strategy for *S* (100 Moves)
		L	R	
A	T	0.95	0.95	100
	B	0.28	1.65	0
B	T	0.68	1.23	60
	B	1.35	0.53	40
C	T	1.02	0.88	90
	B	0.35	1.58	10
D	T	1.12	0.78	75
	B	0.45	1.47	25
E	T	0.60	1.28	50
	B	1.30	0.62	50

All the lines intersect at the point (X′, Y′). X′ is the row player's optimal strategy; in this case 75T–25B. Y′ is the value, V, of the game; in this case 95 points.

The *S* was always the row player. His opponent (*E*) was the column player. The *E*'s choices were obtained as follows: (See Fig. 1). Each strategy of *S* defines an interval (containing V) of possible payoffs to *S*. Each point in this interval corresponds to a strategy for *E*. All points with payoff greater than V (shaded area in Fig. 1) were eliminated so that *S* could never do better than V. The remaining interval shown in the unshaded portion of the graph was divided into 10 equal parts, each part corresponding to the *E* strategies 100–0, 95–5, . . ., 55–45. One of the 10 strategies in the interval determined by *S*'s strategy was selected randomly to give *S* his payoff on that play.

The five 2 × 2 matrix games shown in Table 1 were used. These games were selected according to the following specifications: (*a*) All the games have the same value (V = .95). (*b*) The optimal *S* strategies for the games are: 100–0, 90–10, 75–25, 60–40, 50–50, these numbers referring to the proportions of T and B, respectively. (*c*) The average *S* payoffs for strategies deviating from the optimum are approximately equal for equal absolute deviations and *E*'s most punishing choices (always either 100–0 or 0–100). Graphically this means that the lines representing *E*'s 0–100 and 100–0 strategies have slopes of approximately equal magnitude, but opposite in sign.

The payoffs for the five games were obtained graphically, by drawing the lines *ac* and *bd*, as in Fig. 1, and reading off the intercepts at pairs of perpendiculars 100x units apart; the left perpendicular at x′ (100–0 game), 10 units to the left of x′ (90–10 game), etc. The values were rounded to the nearest hundredth. The slopes of the lines *ac* and *bd* were chosen with two ideas in mind.

In the first place, if the slopes were too small, differences in players' responses would not be differentially rewarded and it would be difficult to locate the optimal strategy by trial and error. On the other hand, if the slopes were large, the range of payoffs would be large, leading both to easier response differentiation and to awkward point-to-money conversion rates.

In the second place, the two lines were chosen with slightly different absolute slopes to avoid obvious symmetries in the 50–50 game and to avoid the easily learned characteristic of equal slopes, i.e., probability of playing a given row is simply the difference between the payoffs in the other row divided by twice the slope. Furthermore, the slope for any line is simply the difference between the column payoffs. Therefore, for equal slopes, the optimal strategy would always be the difference between the payoffs in the opposite row, divided by twice the difference between the payoffs in either column. The values actually used in the experiment reflect a compromise among these conflicting considerations. The slopes of the lines *ac* and *bd* were the same for all five games. Since these slopes determine a reward gradient for responses deviating from optimal strategies, it is clear that the reward gradients were the same for all games. This means that the average reward and range of rewards for responses deviating by a given absolute amount from optimal strategies was the same for all games. The games were obtained empirically rather than analytically partly because of the ease with which the graphic method provided a solution to these problems and partly because two place accuracy was considered more than sufficient for the purposes of this experiment.

The *S*s were told that their point total would be converted to money after the experiment—the more points, the more money—with a conversion factor of about 100 points equal to 1 cent. They could receive up to $1.00 more than the minimum guaranteed ($1.50) for participating in the experiment. They were urged to make all reasonable efforts to maximize their point total for each trial and for the entire series of tasks. The *S*s were, in fact, paid on a linear scale where the lowest point total equaled $1.50 and the highest equaled $2.50.

The same game value was used for all five games to avoid confounding the effect of value with that of the major variable, optimal strategy. The value of 95 points was chosen so that the point-to-money conversion could be made at reasonable cost, i.e., it was felt that a ratio of approximately 100 points to 1 cent would give a return satisfactory to *S*s and tolerable to the experimental budget. The value of the games was not made a rounded figure, say 100 points, in order to avoid making this particular aspect of the procedure too transparent to *S*s.

Instructions.—In the instructions given *S* the following points were emphasized: (*a*) Your ability to make decisions is being tested. The better your decisions, the higher will be your payoff. (*b*) The *E* is single-minded in his determination to keep your winnings as low as possible. He will be informed of your choices but his information will not be complete. He will be told enough, however, so that unwise choices on your part can be punished severely. (*c*) A given payoff will be determined entirely by two factors: your strategy and your opponent's strategy.

The procedure for obtaining the payoffs from the two strategies at each trial was explained carefully to *S*s. One reason was to avoid the misconception that if, for example, the two strategies were T90–B10 and L90–R10, there would be 90TL combinations, 10BR combinations and no BL or TR combinations, when in fact there would be 81TL, 9TR, 9BL, and 1BR combinations.

The procedure was represented to *S*s as an analog of real life situations involving two choices each for two antagonists. The payoff rules of the game as given by the matrix were explained, but no instructions were given involving game theory concepts. Specifically, *S*s were not told that all games had the same value, nor what the value of any game was, nor indeed did they have any explanation of what a game value was. They were told nothing of the possible payoff beyond the information contained in the matrix and the important facts labeled (*b*) immediately above. All this was aimed at creating a situation in which *S*s were confronted by a hostile and knowledgeable opponent. By choosing their optimal strategy *S*s could guarantee themselves a certain number of points each trial. They could never do better than the optimal strategy payoff but they could do considerably worse.

Procedure.—The procedure for each game was as follows: *S* was given a three-page answer sheet with the matrix game reproduced on each page. The alternatives were denoted by capital letters and were changed from game to game and from *S* to *S*. The *S* announced his strategy (frequencies of the T and B alternatives adding to 100). The *E* chose a strategy punishing to *S* as described above. The *E* then announced the strategy he had chosen together with *S*'s payoff. The *S* recorded both strategies and his payoff. This procedure was continued for 50 trials or until criterion was reached. (See Results and Analysis).

Experimental design.—The design, shown in Table 2, was a 5 × 5 latin square replicated four times; each replication had the 5 games, 5 orders of presentation (each with a different *S*), and 5 temporal positions.

TABLE 2

EXPERIMENTAL DESIGN

Group	*N*	Order of Presentation of Games				
		1	2	3	4	5
I	4	A	B	C	D	E
II	4	B	A	D	E	C
III	4	C	E	A	B	D
IV	4	D	C	E	A	B
V	4	E	D	B	C	A

The order of the rows in each game matrix was varied randomly from one *S* to another. Since each game in each of the five order of presentation groups was presented to four different *S*s, two of the four *S*s in each group had the matrix as given in Table 1, and the other two had the rows inverted.

Subjects.—Twenty paid volunteers, undergraduate women students from Connecticut College, served as *S*s.

RESULTS

Two measures were used, each corresponding to a somewhat different aspect of performance. To measure the overall correspondence of the response to the optimal strategy an "integrated error" score was used. This was the average deviation (*AD*) from optimal strategy over all trials for each game. These are shown in Tables 3 and 4. The other measure was trials to achieve a steady-state response. Steady-state was defined as repetition, without deviation, of the same response up to and including the last trial (Trial 50), or *S*'s announced intention to continue in this way. The percentage of *S*s reaching criterion in each of the five positions was 10, 20, 40, 50, and 65, and in each of the five games was 50, 40, 30, 35, and 30 for Games A through E, respectively. All steady-state solutions to the game problems were the "correct" game theory (maximin) solutions, but a steady-state was not always achieved.

Because a steady-state solution was not achieved in many games, the analyses of the criterion scores were performed using a nonpara-

TABLE 3

FREQUENCY DISTRIBUTIONS FOR *AD*s FROM OPTIMAL STRATEGY BY POSITIONS

AD	Position				
	1	2	3	4	5
40+	0	1	0	0	1
30–39.9	4	1	1	1	2
20–29.9	4	7	0	2	1
10–19.9	9	7	11	6	5
0–9.9	3	4	8	11	11
Mean	19.5	19.0	12.0	11.5	13.5

TABLE 4

FREQUENCY DISTRIBUTIONS FOR *AD*s FROM OPTIMAL STRATEGY BY GAMES

AD	Game				
	A	B	C	D	E
40+	1	0	1	0	0
30–39.9	3	1	2	2	1
20–29.9	6	0	2	3	3
10–19.9	6	7	8	7	10
0–9.9	4	12	7	8	6
Mean	20.5	10.0	16.0	14.5	14.5

TABLE 5

ANALYSIS OF VARIANCE OF *AD* SCORES

Source	*df*	*MS*	*F*
Between *Ss*			
Order of presentation	4	109.86	
Replicates (Groups)	3	375.65	1.62
Residual among *Ss*	12	231.84	3.94*
Within *Ss*			
Games	4	268.59	4.56*
Position	4	388.14	6.59*
Residual from L.S.	12		
Residual within *Ss*	60	72 58.91	

* $P < .01$.

metric ranking test, the Friedman two-way analysis of variance on ranks.

Average deviation scores.—The analysis of variance performed on the *AD* scores is shown in Table 5. The results show that performance, using this measure, varied with the person playing the game, the game being played, and the point in the sequence at which the game was played.

Trials-to-criterion scores.—In performing the analysis of variance on ranks, parallel analyses were done on rankings with ties, and on rankings with ties broken by appeal to the *AD* scores. The results, shown in Table 6, indicate that performance, using this measure, varied with the person playing the game and the point in the sequence at which the game was played, but did not vary with the game.

Of the 100 games played by the 20 *Ss*, 37 were solved by 13 *Ss*. Among the 13 *Ss* who solved at least one game the mean number of solutions was 2.8 (median of 3) out of a possible five. In only one case did an *S* fail to solve a game following a solution to an earlier game.

In addition to the above two response measures the point totals collected by each *S* for each game were analyzed to see what effects were reflected by the amount of points received by *Ss*. Although the main concern was centered on what *Ss* did (responses given as strategies) rather than what they received (reinforcement given as number of points received) there is a possibility that the lack of response differentiation might have been mediated by reward constancy, or in simpler terms, that *Ss* failed to seek and find better strategies because the extra effort did not lead to any particular gain in payoff. The analysis of variance on the point totals reveals the same effects as those demonstrated in the response measure. The tau correlation between earning (points) and deviations from optimal strategy, computed over *Ss*, was -0.947 ($P < .001$).

The "how-much" scores of average error and points, and the "how-long" score of trials-to-criterion lead to similar conclusions concerning effects of *Ss* and temporal position. A major discrepancy is found for the five games which differ in their optimal strategies for *S*. Here it is found that *Ss* tend on the average to be further from the optimal strategies on the 100–0 and 90–10 games than on the other three games in the presolution portions of the game, but do not differ significantly in the number of trials taken to achieve a solution.

TABLE 6

FRIEDMAN TWO-WAY ANALYSIS OF VARIANCE ON RANKS OF TRIALS-TO-CRITERION SCORES

Source	*df*	χ^2	
		Ties	No Ties
Games	4	2.4	5.0
Position	4	15.0*	27.7*
Ss	19	36.7*	45.1*

* $P < .01$.

A possible explanation of this result is that the true *AD* for the different games would be expected to be unequal under the reasonable assumption that *S*s are responding randomly over the entire strategy range from 0–100 to 100–0. Thus, differences among mean *AD* could arise as an artifact from a situation where *S*s respond similarly (randomly) in all the games.

To test this further, mean *AD* scores were measured as differences between observed *AD* and expected *AD* under the assumption of random responding.

An *F* test, using the same error term as previously, showed the adjusted *AD*s again to differ significantly ($F = 4.50$ with 4/72 *df*). Although the procedure did not take account of the differences in variability that the random responding hypothesis would predict, the results were judged to be sufficiently clear-cut to justify a rejection of the hypothesis.

These results indicate that the hypothesis of random responding does not account for the observed differences in mean *AD* among the games.

The fact that the different games do not differ significantly in trials to criterion in no way conflicts with the significant findings above, using *AD*. There is no reason to believe that the two measures (*AD* and trials-to-criterion) are strongly dependent. The results suggest in fact that they are not.

The two tests above show that *S*s responses (*AD*) differed significantly among the five games, and that these differences would not be accounted for by a simple hypothesis of random responding. A hypothesis that is not incompatible with these results is that *S*'s *AD* in each game is a constant proportion of the *AD* expected under random responding. An analysis of variance was performed on mean *AD* taken as a proportion of the expected random *AD*. Results of this analysis of variance for *S*s, order of presentation, replicates, and temporal position were similar to those of the previous analysis, but the variability for games was now found to be not significant ($F = 2.26$, with 4/72 *df*, $P > .05$)

These results indicate that, in answer to the second question posed in the introduction: (*a*) In general, the more extreme (differing from 50–50) the optimal strategy, the more players deviate from the optimum. (*b*) In general, the more extreme the optimal strategy, the more improvement players show over random performance. (*c*) The amount that players deviate from optimal strategy, measured as a proportion of the amount they would have deviated if they were responding at random, is not significantly different from game to game.

Discussion

The results clearly demonstrate some major effects in this game playing situation. A majority of naive *S*s learned to find the game-theory solution through the experience of playing the games against an unfriendly but variable opponent. There were, however, significant differences among *S*s in the degree to which the games were mastered. In psychological terms, these *S*s demonstrated the ability to differentiate one of a large number of possible responses (101 strategies). The conditions under which *S*s learned in this experiment differed in some ways from previous experiments on response differentiation: All responses were rewarded, the response corresponding to optimal strategy was rewarded at least as much as and no less than any other response. The *S*'s task was to pick from among a set of rewarded responses that response which was associated with the maximum reward. This

type of learning has received little experimental attention to date.

An unevaluated factor was the degree to which *S*s were able and willing to utilize the cues given by the payoff matrix. As the problem was presented, *S*s could achieve a solution without referring to the matrix. Theoretically, the payoff matrix was a sufficient but not a necessary cue for obtaining the optimal strategy. The fact that no *S* reached criterion at the beginning of any game, suggests that *S*s relied more heavily upon game playing experience than upon the matrix cue.

Summary

Strategies chosen by players in a real game situation were compared to those given by game theory in order to determine how readily, if at all, naive persons adopt "optimal" rational solutions. The players were forced to choose a set of moves, i.e., a strategy. They received monetary rewards, the average amount varying according to how close their strategy approximated the game theoretical strategy. Each player participated in several games. The optimal strategies differed from game to game, but the expected payoff was held constant. The games were presented in different sequences to the different players in order to determine the effect, if any, of the order of play. The results indicate that individuals differ markedly in their ability or willingness to adopt a game theoretic strategy. However, a majority of *S*s did achieve a game solution in at least one of the five games played. The degree to which the players approximated the game theory solution was a function of the optimal strategy associated with each game.

References

Atkinson, R. C., & Suppes, P. An analysis of two-person game situations in terms of statistical learning theory. *J. exp. Psychol.*, 1958, **55**, 369–378.

Atkinson, R. C., & Suppes, P. Application of a Markov model to two-person noncooperative games. In R. R. Bush & W. K. Estes (Eds.), *Studies in mathematical learning theory.* Stanford, Calif.: Stanford Univer. Press, 1959. Ch. 3.

Lieberman, B. Human behavior in a strictly determined 3 × 3 matrix game. *Amer. Psychologist*, 1959, **14**, 418. (Abstract)

Simon, H. A. A comparison of game theory and learning theory. *Psychometrika*, 1956, **21**, 267–272.

(Received April 6, 1960)

Journal of Experimental Psychology
1961, Vol. 61, No. 6, 469–473

STIMULUS GENERALIZATION AND AGGRESSIVE VERBAL STIMULI [1]

ARNOLD H. BUSS

University of Pittsburgh

This report describes a new method for investigating stimulus generalization with meaningful verbal materials, words that connote aggression; since the stimuli are words, the appropriate term is *mediated* stimulus generalization. It is necessary first to arrange the words on what Thurstone has called a *subjective metric* in order to obtain a stimulus dimension, in this instance a dimension of intensity of aggression. A response may then be conditioned to one part of the dimension, and the spread of this response to other parts of the dimension can be determined.

METHOD

Scaling.—A list of 185 words connoting various degrees of aggression was presented in alphabetical order to 60 college men and 78 college women. Each word was rated for aggression on a 9-point scale, with 1 representing minimal aggression and 9 maximal aggression. After words with excessive variability in ratings were discarded, there remained 146 words, and the median was used as the scale value for each word.[2] The following stimuli represent the range of the intensity dimension:

contest	1.9	bitterness	5.5
controversy	2.4	insulting	6.0
blunt	3.0	rage	6.5
ill-humored	3.5	assailant	7.0
pest	4.0	hatred	7.4
grouch	4.5	bomb	8.0
disapproval	5.0	annihilation	8.6

These are the scale values for the male raters; separate scale values were derived for women, and in general there was close agreement between the sexes. In obtaining generalization gradients the men's scale values were used for male *Ss* and the women's scale values for female *Ss*. None of the students who rated the words was used as an *S* in the experiment.

Design and procedure.—The design and procedure for the first four groups will be described first. There were two series of trials, one for training and one for generalization. The training series consisted of 20 trials, with one stimulus word per trial. These were the 20 most aggressive stimuli that had been scaled (kill, mutilate, etc.). The words were presented singly on a memory drum for 4 sec. each.

The *Ss* were instructed to make one of two responses to each word, and the groups differed in the response that was reinforced. The first two groups were instructed to read the word aloud or to remain silent. Group I was reinforced (the *E* said *Right*) for reading aloud and negatively reinforced (the *E* said *Wrong*) for remaining silent. Group II was reinforced for remaining silent and negatively reinforced for reading aloud.

Groups III and IV were instructed to shout each word or to whisper it. Before the experiment began *E* insured that each *S* would shout loudly and whisper softly, and during the experiment proper *E* could easily distinguish between the shout and whisper responses. Group III was reinforced for shouting, negatively reinforced for whispering. Group IV was reinforced for whispering, negatively reinforced for shouting.

The generalization series for these four groups consisted of the 126 stimuli remaining after the 20 most aggressive stimuli (the training stimuli) were removed. Each word constituted a trial; there were 126 generalization trials, each with a different word. These first four groups were thus trained on the most intensely aggressive stimuli and generalized from these intense stimuli to milder, less aggressive stimuli in the generalization series.

The second four groups generalized in the

[1] This investigation was supported in part by Research Grant M-2861 from the National Institute of Mental Health, United States Public Health Service. John Neisworth and Mary Alice Geiselhart helped in collection and analysis of the data.

[2] A detailed account of the scaling procedures and a complete list of the men's and women's scale values may be found in Buss (in press, Table 7.1).

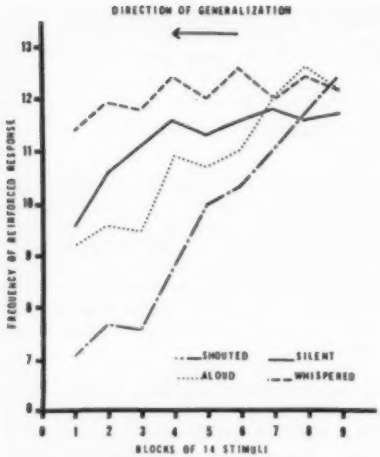

FIG. 1. Generalization gradients for Groups I to IV.

opposite direction: from mildly aggressive words to intensely aggressive words. Their training stimuli were the 20 least aggressive words (controversy, blunt, etc.). The procedure was identical to that for the first four groups, and the second four groups paralleled the first four in the response reinforced: Group V (Aloud), Group VI (Silent), Group VII (Shout), and Group VIII (Whisper). Again the generalization series consisted of the 126 remaining words after the training stimuli were removed.

At the end of the training series each *S* was instructed to continue responding as he had been but that *E* would no longer tell him whether he was correct or not. Then the generalization series began, each stimulus being presented for 4 sec. The order of presentation was a random one that was the same for all *Ss*.

Subjects.—The *Ss* were college men and women, recruited from introductory psychology classes. They were randomly assigned to the eight experimental groups with the restriction that the number of men and women in each group be proportionate. There were 40 *Ss* in Groups I and II and 20 each in the remaining six groups, making a total of 200 *Ss*.

RESULTS

The 126 stimuli were grouped on the basis of intensity into nine blocks of 14 stimuli each, with Block 1 representing low intensity and Block 9 high intensity. The first four groups received training at the intense end of the aggression dimension and generalized to less intense stimuli. Their (mediated) stimulus generalization gradients are presented in Fig. 1. Inspection of this figure reveals that the curves differ in slope, and in order of increasingly steep gradients they are: Whisper, Silent, Aloud, and Shout. In order to test the significance of these differences an analysis of variance was performed, the results of which are presented in Table 1. The Groups × Stimuli interaction is the statistic relevant to slope differences, and this interaction is significant at the .001 level. In order to determine the statistical reliability of slope differences between adjacent gradients, three separate analyses of variance were performed: Whisper vs. Silent, Silent vs. Aloud, and Aloud vs. Shout. In these analyses the Groups × Stimuli interactions were significant at the .01, .05, and .001 levels, respectively. Thus, each gradient differed significantly in slope from its neighbor.

The four different responses used in this study may be aligned on a dimension of activity: silent, whisper, aloud, and shout, in order of increasing activity or intensity. This alignment differs from the order of steep-

TABLE 1

ANALYSIS OF VARIANCE FOR GROUPS I TO IV

Source	*df*	*MS*	*F*
Groups	3	182.63	1.4
Error (between)	119	130.01	
Stimuli (blocks of 14)	8	140.53	94.3*
Groups × Stimuli	24	10.97	7.4*
Error (within)	925	1.49	
Total	1079		

* $P < .001$.

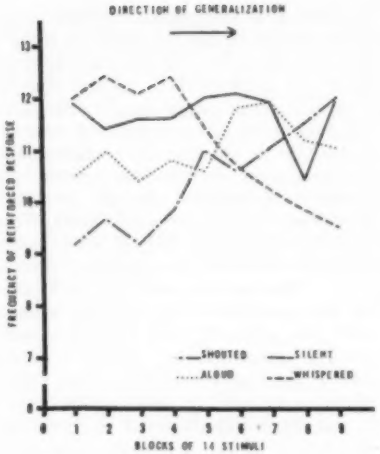

FIG. 2. Generalization gradients for Groups V to VIII.

ness of the gradients only in the position of the Silent gradient; if the Silent gradient is omitted, Fig. 1 reveals that in order of increasing steepness the gradients are: Whisper; Aloud, and Shout. This sequence matches the sequence of activity or intensity level of the response.

The second four groups (V to VIII) received training at the mild end of the aggression dimension and generalized to more intense stimuli; their gradients are presented in Fig. 2. These gradients are considerably flatter than their counterparts in Fig. 1, and there is much overlap between gradients. The intertwining of the curves makes it difficult to see differences in the slopes, but it is clear that the Whisper gradient has the steepest slope and that the Shout gradient is *inverse* (fewer shout responses near the point of training than farther away). Differences between the gradients were tested by an analysis of variance, which is presented in Table 2. The significant Groups × Stimuli interaction indicates that the differences in slope are statistically reliable. Analyses of variance comparing two groups at a time reveal that, except for the difference between the Silent and Aloud groups, all differences between adjacent gradients are significant.

Figure 2 may be examined concerning the relationship between activity of the reinforced response and the steepness of the gradient. Again the pattern is clearer if the Silent group is omitted. The Whisper gradient is clearly steeper than the Aloud gradient, and the Shout gradient is not merely flat but rises instead of falling away from the point of training. The order of steepness away from the point of training is: Whisper, Aloud, and Shout. Thus, when the response generalizes from mild stimuli to more intense stimuli, the steepness of the gradient is inversely related to the activity of the response. This is exactly opposite to what occurs when the response generalizes from intense stimuli to milder stimuli, where the steepness of the gradient is *directly* related to the activity of the response.

DISCUSSION

The results may be accounted for in terms of a "matching principle" and direction of generalization. The matching principle states that there is an

TABLE 2

ANALYSIS OF VARIANCE FOR GROUPS V TO VIII

Source	*df*	*MS*	*F*
Groups	3	46.38	.7
Error (between)	76	64.30	
Stimuli (blocks of 14)	8	3.41	1.1
Groups × Stimuli	24	18.80	5.9*
Error (within)	608	3.17	
Total	719		

* $P < .001$.

approximate matching of the intensity (activity level) of the response to the intensity of the stimulus. Thus in Groups III and IV there should be a tendency to shout intensely aggressive words and to whisper mildly aggressive words. Since these groups generalized from intensely aggressive to mildly aggressive stimuli, this tendency should steepen the generalization gradient of Group III (Shout) and flatten the gradient of Group IV (Whisper). The gradient of Group I (Aloud) should be flattened slightly because the Aloud response is intermediate between shouting and whispering. (The Silent group, II, will be discussed below.) Figure 1 reveals that the gradients are in accord with the prediction made on the basis of the matching principle: Shout, Aloud, and Whisper, in decreasing order of steepness.

In Groups VII and VIII there should be a strong tendency to whisper mildly aggressive words and to shout intensely aggressive words. Since these groups generalized from mildly aggressive to intensely aggressive words, this tendency should steepen the Group VIII (Whisper) gradient and flatten (or reverse) the Group VII (Shout) gradient; again the Aloud gradient should be intermediate. Figure 2 indicates that these expectations are fulfilled. The matching principle applies particularly to the inverse gradient of the Shout group (VII), the tendency to shout intensely aggressive words being the only explanation for an *inverse* generalization gradient in the present context.

Concerning the Silent response, it is suggested that "no response" does not belong on a continuum of response intensity or activity. The *S* must *suppress* a vocal response, and the effects of inhibiting a response on stimulus generalization are not known.

A comparison of the mediated stimulus generalization gradients in the two figures reveals that those in Fig. 1 are somewhat steeper than those in Fig. 2. Note that the Stimulus *F* is significant for the first four groups but not for the second four groups (see Tables 1 and 2), indicating a statistically reliable slope when the generalization is from intense to mild stimuli but not when the generalization is from mild to intense stimuli. This difference in slope may be attributed to the stimulus intensity dynamism (Hull, 1943), but another explanation is perhaps more reasonable. When the intensely aggressive stimuli (murder, mutilate) were presented during the training series, they helped *S* to formulate some notion of the aggression dimension. The process was also facilitated by the contrast of the training stimuli with the generally less intense stimuli of the generalization series. When mildly aggressive stimuli (controversy, blunt) were used in the training series, it was more difficult for *S* to formulate the nature of the dimension, and later there was less contrast with the generalization stimuli. Therefore it is not surprising that *S*s trained on mildly aggressive stimuli had flatter gradients than those trained on intensely aggressive stimuli.

A final word should be added about the method employed in this study. It is a new method of studying stimulus generalization in that it employs meaningful verbal stimuli that have been scaled along a personality dimension and also in that *S* was trained with a *range* of stimuli rather than with a single stimulus. Other aspects of the paradigm, however, are similar to that used in other stimulus generalization studies, and the mediated generalization gradients are in many respects comparable to those derived from such studies. In light of these considerations and the results that emerged, it would seem that the method may appropriately be extended to the study of stimulus generalization with other kinds of meaningful materials.

Summary

A method for studying stimulus generalization with verbal stimuli was described. Words were scaled for aggression and presented singly on a memory drum. For four groups the training stimuli were the most aggressive words, and for four groups the

training stimuli were the least aggressive words. In each set of four groups, one group was reinforced for shouting, one for reading aloud, one for whispering, and one for remaining silent. After training, all groups received a generalization series consisting of words varying in intensity of aggressive connotation.

Evidence was found for operation of a matching principle, whereby the intensity of the verbal response tends to vary with the intensity of the verbal stimulus. Generalizing from intense to mild stimuli yielded steeper (mediated) generalization gradients than generalizing from mild to intense stimuli. The method may prove useful for investigating stimulus generalization with meaningful material.

References

Buss, A. H. *The psychology of aggression.* New York: Wiley, in press.

Hull, C. L. *Principles of behavior.* New York: Appleton-Century-Crofts, 1943.

(Received April 7, 1960)

Journal of Experimental Psychology
1961, Vol. 61, No. 6, 474–479

EFFECT OF VOLUME OF REINFORCEMENT AND NUMBER OF CONSUMMATORY RESPONSES ON LICKING AND RUNNING BEHAVIOR[1]

HARRY L. SNYDER AND STEWART H. HULSE

Johns Hopkins University

In this experiment, the total volume (cc.) of water reinforcement and the total number of consummatory responses (licks on a drinking tube) were independently varied in a factorial design. Licking behavior and running behavior were analyzed as a function of these variables. Data of this sort are particularly relevant to some typical studies concerned with amount of reinforcement which have confounded number of consummatory responses with total weight or volume of the reinforcing substance (e.g., Crespi, 1942; Hulse, 1958; Zeaman, 1949).

In addition, performance was analyzed as a function of *volume per lick*, a variable which has recently been shown to relate in a systematic fashion to the rate of emission of the consummatory response (Hulse, Snyder, & Bacon, 1960).

METHOD

Subjects and Apparatus

The *Ss* were 45 experimentally naive male albino rats of the Wistar strain which were obtained from Albino Farms, Red Bank, New Jersey. They ranged in age from 70 to 90 days at the start of taming. Five *Ss* were randomly assigned to each of nine experimental groups. Three *Ss* died during the course of the experiment. These were treated statistically as missing data since it was not possible to replace them with comparable *Ss*.

The *Ss* were run in an L-shaped, enclosed, wooden runway whose total length was 74 in. The startbox was 7 in. long and 5½ in. wide. The running alley was 54 in. long and 4 in. wide, and the goalbox was 20 in. long and 4 in. wide. The runway had a height of 4 in. and was covered throughout by a hinged Plexiglas lid. The goalbox floor was covered with brass shim stock. All wooden parts of the runway were painted flat black.

One metal guillotine door separated the startbox from the running alley, and another separated the running alley from the goalbox. Two photoelectric cells were located in the running alley ½ in. above the floor; these were used to facilitate the timing of running. The first photoelectric cell was placed 2 in. past the startbox door, and the second 2½ in. past the goalbox door.

A 3½-in. square piece of Plexiglas was located in the wall at the far end of the goalbox. A drinking tube was centered in a ¾-in. × ⅜-in. vertical slot cut in the Plexiglas. The drinking tube consisted of a piece of ¼-in. brass tubing which contained a 2-mm. fluid hole.

With each lick on the drinking tube, a pumping system was activated which instantly delivered a specified volume of water at the tip of the drinking tube. The pumping system, which is described in more detail elsewhere (Hulse, 1960), consisted of an infusion pump operated by an electronic "drinkometer" circuit. The "drinkometer" circuit also operated a counter and a Brush Development Co. recorder, Type BL202, which provided records of licking behavior. The recorder's paper speed was 2.5 cm. per sec.

Experimental Design

Three levels of total volume of water reinforcement (2.65, 1.50, and 0.75 cc.) and three levels of total licks (295, 450, and 675 licks) were combined factorially to produce nine experimental groups. Each group received a different fixed *volume per lick*, so that the Total Volume and Total Lick requirements were satisfied. For example, those *Ss* receiving 2.65 cc. of water distributed over 295 licks were given .0090 cc. per lick.

[1] This research, which represents the MA thesis of the senior author, was supported by a National Science Foundation research grant (G8712) to the junior author.

In this manner, the values of volume per lick ranged from .0011 to .0090 cc. per lick.

Procedure

Taming.—On each of the 5 days of taming, *Ss* were handled freely and permitted to drink water from a dish for 10 min. This was the only water available during taming. Ad libitum food was available in the individual living cages throughout the experiment.

Training.—Following taming, *Ss* were given 34 training trials, one trial per day. The order of running was randomized daily for all *Ss*. On each trial, *S* was placed in the startbox with the door lowered. After a delay of 2–3 sec., the starting door was raised. Retracing was prevented by lowering each door immediately after *S* had passed. When the allotted number of licks had been made upon the drinking tube, *S* was removed from the goalbox and placed into a wooden retaining cage. After 15 min., *S* was given a dish containing enough water to bring its daily ration to 14 cc. The *S* was then returned to its home cage.

On Days 1 and 2 of training the drinking tube was centered in the slot so that it extended ¼ in. into the goalbox. On Day 3 the tube was recessed so that it was flush with the inner edge of the slot, and on Day 4 and thereafter the tube was located ¼ in. behind the inner edge. This recession of the tube prevented *S*'s enclosing the tube in its mouth and assured discrete licking responses.

Extinction.—Beginning with Trial 35 and continuing for 13 days, no water was delivered through the drinking tube, although *Ss* were permitted to lick upon the dry tube. The *Ss* were confined for 20 sec. in the goalbox on each extinction trial. All other conditions remained the same as during training.

Measures of Performance

Licking was measured in two ways during training. First, *average rate of licking* was determined to provide an overall estimate of rate of responding during training. This measure was obtained for each *S* for each trial and consisted of the total licks divided by the time required to emit the total licks. Second, *momentary rate of licking* was calculated for short samples of licking taken on Days 28, 31, and 34 of training. This provided a measure of rate of licking which excluded time that *S* spent doing things other than lick. Thus, 10-sec. Brush kymograph records were obtained for the 675-lick *Ss* at 5, 30, 60, and 90 sec. following the first lick, and for the 295 and 450-lick *Ss* at 5 sec. and 30 sec. following the first lick. For each record, intervals between licks which were longer than twice the average interlick interval for that particular record were subtracted from the total 10-sec. duration of the record. The number of licks appearing on the record was then divided by the remaining time to give momentary rate.

During extinction, licking behavior was measured by recording the number of licks emitted during the 20 sec. of goalbox confinement.

Latency of running was defined as the time which elapsed from the moment the starting door was raised until the moment *S* broke the first photocell beam. *Time of running* was defined as the time which elapsed during *S*'s passage between the two photocells. Both latency of running and time of running were automatically recorded on Standard Electric timers. Time of running was converted to a reciprocal and multiplied by 100 to give *speed of running.*

Results

Licking behavior.—The average rate of licking is not affected by either Total Volume or Total Licks, per se (Table 1). However, the data show a significant Volume × Licks interaction. This interaction variance is derived from unique combinations of Total Volume and Total Licks. In this particular design, such combinations may be considered equivalent to different *volumes per lick.* As shown in Fig. 1, the average rate of licking is a nonmonotonic function of volume per lick.

Momentary rates of licking could not be obtained for all *Ss* after 60 sec. and 90 sec. of licking because many *Ss* completed their allotted number of licks before that time. Consequently, it was necessary to examine the momentary rates in separate analyses of variance, as shown in Table 1. The first analysis treated the momentary rates for all *Ss* for 5- and 30-sec. intervals, and the second analysis treated the 5-, 30-, 60-, and 90-sec. intervals for the 675-lick *Ss*.

TABLE 1

ANALYSES OF VARIANCE FOR AVERAGE AND MOMENTARY RATES OF LICKING DURING ACQUISITION

Source	Average Rate		Momentary Rate (5 and 30 sec.)		Momentary Rate (5, 30, 60, and 90 sec.)	
	df	*MS*	*df*	*MS*	*df*	*MS*
Ss	(42)[a]	—	(42)	—	(13)	—
Volume (V)	2	1.148	2	15.592**	2	6.264*
Licks (L)	2	.473	2	.172	—	—
V × L	4	3.233*	4	1.495	—	—
Between *Ss* within V, L	34	1.120	34	.818	11	1.291
Trials	(2)	—	(5)	1.854**	(11)	.647**
Days (D)	2	.084	2	.180*	2	.043
Time (T)	—	—	1	8.881**	3	2.310**
D × T	—	—	2	.016	6	.016
Ss × Trials[b]	(84)	.065	(210)	.051	(143)	.033
D × *Ss*	68	.064	68	.088	22	.081
T × *Ss*	—	—	34	.071	33	.057
T × D × *Ss*	—	—	68	.008	66	.003
Pooled error[c]	—	—	170	.053	121	.032

[a] Appropriate corrections for missing data have been made and the error *df* altered accordingly (Lindquist, 1956, p. 148).
[b] Of the *Ss* × Trials interactions, only V × T was significant ($F = 2.75$, $df = 6/121$, $P < .05$) for the 5-90 sec. momentary rate analysis.
[c] Under the assumption of no intrinsic interactions between *Ss* and any of the variables, the pooled error term was used in order to obtain a more stable error estimate.
* $P < .05$.
** $P < .001$.

The momentary rate of licking increases with Total Volume, but decreases with time since the first lick (Fig. 2 and 3). This decrease with time is significantly greater for those *Ss* receiving larger Total Volumes ($P < .05$).

While the average rate of licking is nonmonotonically related to volume per lick, Fig. 3 shows that the momentary rate of licking is an increasing monotonic function of volume per lick. In this instance, because of the monotonic nature of the function and the confounding of volume per lick with total volume and total licks, the effect of volume per lick is not indicated by a significant Volume × Lick interaction. However, a rank-order correlation between volume per lick and the mean momentary rate of each group yielded a rho of +.82 ($P < .01$).

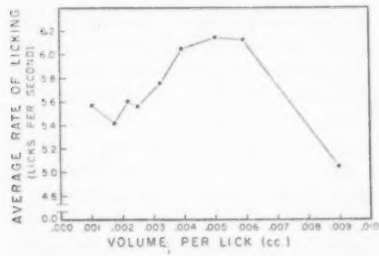

FIG. 1. Average rate of licking during acquisition as a function of volume per lick.

An analysis of the mean licks during extinction indicated no significant variation due to either total volume, total licks or, once again, the Volume × Licks interaction. However, a rank-order correlation of −.83 ($P < .01$) was obtained between the mean number of licks for each of the nine groups and volume per lick. Consequently, volume per lick

appears to be the main variable correlated with resistance to extinction of the licking response.

Running behavior.—Analyses of the mean running speeds and the mean running latencies for Days 21–35 of training and Days 36–47 of extinction failed to show any significant differences due to either total volume, total licks, or their interaction ($P > .05$). Asymptotic performance levels during training were assumed since there was no significant increase in either of the running measures between Days 21 and 35 as indicated by these analyses ($P > .05$).

Discussion

Our failure to find any consistent relationship between running speeds and either total volume, total licks, volume per lick, or rate of licking is of some interest since theorists such as Spence (1956) and Sheffield, Roby, and Campbell (1954) have argued that number of consummatory responses and "vigor" of consummatory behavior ought to be positively correlated with running speed in the alley. While our results may raise some questions concerning the generality of these theories, no difference data can never, of course, be conclusive. Quite possibly our design did not adequately cover the low end of the Total Licks or Total Volume variables. Since this experiment was completed Goodrich (1960) has reported small but significant differences in running speeds for *Ss* allowed only 5 as compared with 40 or 145 licks in the goalbox.

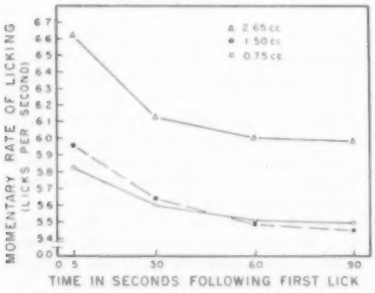

Fig. 2. Momentary rate of licking during acquisition as a function of time since the first lick for three different total volumes of reinforcement.

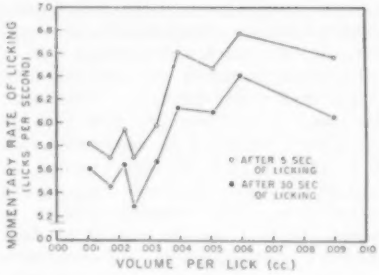

Fig. 3. Momentary rate of licking during acquisition after 5 and 30 sec. of licking as a function of volume per lick.

Although final interpretation of the running data must remain for the future, there is no question that the licking data per se are consistent, particularly with respect to volume per lick. The average and momentary rates of licking form essentially parallel functions between .001 and .005 cc. per lick, with the shapes of the functions becoming quite different for larger volumes per lick. The momentary rate remains approximately at asymptote for values above .005 cc. The rat, when he licks, licks faster the larger the volume per lick. The average rate of licking, on the other hand, *decreases* for values above .005 cc. per lick. The rat pauses more frequently between bursts of licking the larger the volume per lick.

The data suggest that licking behavior varies with volume per lick in much the same way that the rate of bar-pressing varies with different concentrations of either sucrose or saccharine. Thus, the *average* rate of licking is nonmonotonically related to volume per lick just as the rate of bar pressing is nonmonotonically related to the concentration of either sucrose (Guttman, 1953) or saccharine (Cockrell, 1952).

In both cases as amount of reinforcement (defined either volumetrically or qualitatively) increases, the average rate of responding increases to some maximal value and then decreases. Similarly, increases in volume per lick produce monotonic increases in *momentary* rate just as increases in concentration of sucrose produce monotonic increases in initial rate of bar pressing (Collier & Siskel, 1959). As the intake of fluid during a session goes up, however, both rate of bar pressing and momentary rate of licking appear to show the effects of postingestive factors, such as gut-loading or changes in taste sensitivity, which apparently depress rate of responding (Collier & Siskel, 1959). Thus, rate of bar pressing eventually decreases as a function of concentration of sucrose and the frequency with which reinforcements are administered during the experimental session. In a congruent manner, momentary rate of licking decreases as a function of two factors which affect the extent of fluid intake: time since the beginning of licking, and Total Volume or, equivalently, volume per lick (Fig. 2).

The very high, significant, negative correlation between volume per lick and the mean number of licks emitted during extinction is another noteworthy aspect of the licking data. The inverse relationship is not characteristic of experiments involving bar pressing (e.g., Guttman, 1953), but is consistent with the results of some other experiments involving licking or running behavior (Armus, 1959; Hulse, 1958; Hulse et al., 1960).

Finally, we must note that the three parameters of reinforcement used in this study must, by their very nature, remain at least partially confounded. Other things equal, increases in Total Volume or decreases in Total Licks are necessarily accompanied by increases in volume per lick. While in no way can all three of these variables be made orthogonal, it would appear that volume per lick is most consistently and meaningfully related to the differences in licking behavior observed here and elsewhere (Hulse et al., 1960).

Summary

This experiment investigated the effects of total volume of water reinforcement (0.75, 1.50, and 2.65 cc.), total number of licks on a drinking tube (295, 450, and 675 licks), and volume per lick (.0011 to .0090 cc. per lick) upon the running and licking behavior of albino rats. Thirty-five training trials and 12 extinction trials were given, one trial per day, in a 74-in. L-shaped alley.

The training results show that two different measures of licking behavior varied consistently with volume per lick, but not with either total volume or total licks. One measure, average rate of licking, varied nonmonotonically with volume per lick, while the other measure, momentary rate of licking, increased monotonically with volume per lick. Momentary rate was shown to decrease as licking continued within a reinforcement session ($P < .001$). Moreover, this decrease was greater for larger volumes per lick ($P < .05$).

The extinction results show that extinction of the licking response, as measured by the number of licks made upon the dry drinking tube, was slower for small as compared with large volumes per lick ($P < .01$).

Running behavior did not vary significantly with any of the reinforcement conditions during either training or extinction. The possible implications of this observation for certain theories of reinforcement were discussed.

References

Armus, H. L. Effect of magnitude of reinforcement on acquisition and extinction of a running response. *J. exp. Psychol.*, 1959, **58**, 61–63.

Cockrell, J. T. Operant behavior of white rats in relation to the concentration of a nonnutritive sweet substance used as a reinforcement. Unpublished doctoral dissertation, Indiana University, 1952.

Collier, G., & Siskel, M., Jr. Performance as a joint function of amount of reinforcement and inter-reinforcement interval. *J. exp. Psychol.*, 1959, **57**, 115–120.

Crespi, L. P. Quantitative variation of incentive and performance in the white rat. *Amer. J. Psychol.*, 1942, **55**, 467–517.

GOODRICH, K. P. Running speed and drinking rate as functions of sucrose concentration and amount of consummatory activity. *J. comp. physiol. Psychol.*, 1960, **53**, 245–250.

GUTTMAN, N. Operant conditioning, extinction, and periodic reinforcement in relation to concentration of sucrose used as a reinforcing agent. *J. exp. Psychol.*, 1953, **46**, 213–224.

HULSE, S. H., JR. Amount and percentage of reinforcement and duration of goal confinement in conditioning and extinction. *J. exp. Psychol.*, 1958, **56**, 48–57.

HULSE, S. H., JR. A precision liquid feeding system controlled by licking behavior. *J. exp. Anal. Behav.*, 1960, **3**, 1–3.

HULSE, S. H., SNYDER, H. L., & BACON, W. E. Instrumental licking behavior as a function of schedule, volume, and concentration of a saccharine reinforcer. *J. exp. Psychol.*, 1960, **60**, 359–364.

LINDQUIST, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1956.

SHEFFIELD, F. D., ROBY, T. B., & CAMPBELL, B. A. Drive reduction versus consummatory behavior as determinants of reinforcement. *J. comp. physiol. Psychol.*, 1954, **47**, 349–354.

SPENCE, K. W. *Behavior theory and conditioning.* New Haven: Yale Univer. Press, 1956.

ZEAMAN, D. Response latency as a function of amount of reinforcement. *J. exp. Psychol.*, 1949, **39**, 466–483.

(Received April 8, 1960)

Journal of Experimental Psychology
1961, Vol. 61, No. 6, 480–488

PERSISTENCE OF A RESPONSE IN THE APPARENT ABSENCE OF MOTIVATION [1]

HALBERT B. ROBINSON [2]

Stanford University

A long-standing controversy among behavior theorists concerns the possibility that responses may persist in the apparent absence of the motive under which they were acquired. This problem was emphasized a few decades ago by such writers as Woodworth (1918) and Allport (1937), and has since occupied the attention of Hull (1952), McClelland (1942), and others. Learning theorists today seem to agree that this notion (the concept of "functional autonomy") is neither necessary nor valid. Theorists who invoke the concept of drive to explain the learning of an instrumental response reject the functional autonomy concept because they assume the continued importance of some drive state in the persistence of this response. Theorists who do not invoke a drive concept to explain the original acquisition of these responses find the notion unnecessary because they require no drive construct to explain persisting responses.

It is clear that there can be no experimental test of this theoretical question unless the measure of the strength of the "drive" component is separate from the measure of the strength of the "habit" component. Investigators have usually been forced by the design of their experiments to utilize the same measures of the single overt response under consideration as indices of both "drive strength" and "habit strength." The present experiment utilized a situation in which independent measures could be made of drive strength and response strength. This was accomplished by investigating the possible change in performance of a secondary avoidance response after extinction of the primary avoidance (fear) response on which it was based.

The main conjecture can be stated as follows: if a secondary avoidance response is sustained by the drive (or some residue of it) upon which the primary avoidance response was based, then extinction of the primary avoidance response should lead to extinction of the secondary avoidance response.

This follows because a primary avoidance response is based on a drive such as fear. If the response is extinguished, it must be supposed that fear as a drive is no longer active.

If, however, the secondary avoidance response, once established, is independent of the drive related to the primary avoidance response, extinction of the primary avoidance response should leave the secondary response intact. The principal question to be investigated was, then, whether or not the secondary avoidance response would remain intact

[1] This report represents a portion of a doctoral dissertation submitted in 1957 to the Department of Psychology and the Committee on Graduate Study of Stanford University. The author wishes to thank the members of his committee, E. R. Hilgard (Chairman), Douglas H. Lawrence, C. Leland Winder, and William S. Verplanck for their continued interest and encouragement. John Eaton spent innumerable hours assisting in building the equipment and in running animals; Robert Pierce, Leonard Hommel, and Rosalee Tucker also ran animals and tabulated data.

[2] Now at the University of North Carolina.

after the primary avoidance response was extinguished.

METHOD

Subjects

Eighty experimentally naive albino rats were used in the experiment. Of these, 22 males and 22 females, 105 to 109 days old at the beginning of preliminary training, were descendents of a Palo Alto, California strain derived from Wistar stock; 19 males and 17 females, 101 to 112 days old at the beginning of preliminary training, were from the animal colony of the Stanford University Psychology Department which has animals from the Slonaker strain.

Apparatus

Four modified Miller-Mowrer shuttle boxes (Miller, 1948) were used in the experiment. This apparatus consists of two identical compartments with a passageway between. Each compartment, which has as its floor a grid of metal bars, can be separately electrified. The *S* learns to run to the opposite compartment when shock occurs or is threatened in the first compartment. When the shock (or threat) develops in the second compartment, *S* returns to the first.

The stimuli consisted of the unconditioned stimulus (UCS), electric shock through the grid, and the conditioned stimuli (CS), a light, a buzzer, and the stimuli produced by the motion and sound of the door. The shock was produced by a 4500-v., .5-ma. step-up transformer, which was controlled by a variac. The light consisted of a 7-w. bulb located at the end of each compartment. The buzzer was mounted on a board beneath the door and thus directly between the two compartments. The guillotine door, which was controlled by a solenoid, made a relatively loud clapping sound as it was activated. Responses and the onset and termination of stimuli were automatically recorded on an Esterline-Angus recorder. The apparatus, which was electrically controlled and completely automatic, has been described in detail (Robinson, 1957, pp. 18–27).

A lever could be introduced into either of the compartments, permitting control by *S* over the CS. These stimuli (light, buzzer, door) were activated and continued until *S* depressed the lever which kept off the CS until the lever was released. A barrier in front of the door prevented the running response during acquisition of the lever-pressing response.

Procedure

The experiment used a relatively complex procedure, involving a number of steps. The basic experimental plan involved first, teaching *S*s to avoid shock by running to the opposite compartment at the onset of simultaneous CSs; second, teaching them to press a lever to avoid the now aversive CSs without shock; third, extinction of the primary running response to its free-operant level; and fourth, measuring the effect of this extinction on the strength of the secondary response. The third and fourth steps were repeated. The six stages of the experiment were as follows:

Stage I: Acquisition of the primary avoidance response.—Two groups of *S*s were taught the primary avoidance response, running. Group I consisted of 16 males and 16 females; Group II was composed of 13 males and 11 females. The procedure for the acquisition of the primary running response was as follows:

(*a*) Step A: Pre-experimental adaptation to the apparatus. The *S* was placed in one compartment of the shuttle box and the CSs were activated. The UCS was not used. The CSs continued for a 15-sec. period, or until *S* deactivated them by moving from one compartment to the other. There was a 15-sec. "rest" between the termination of the CSs and the onset of the CSs at the beginning of a new trial. Each *S* was placed in this situation for 5 min. on each of three successive days.

(*b*) Step B: Evocation of the primary running response. The day following Step A, *S*s were placed in one compartment and the CSs activated. Five seconds after the onset of the CSs, the UCS (electric shock) was turned on for the first time. The CSs and UCS continued until *S* moved to the opposite compartment. Following 1 min. of rest, the stimuli were introduced into the second compartment. The equipment was controlled manually during these breaking-in trials. The shock level was just subtetanizing and was individually determined for each *S*. Each *S* underwent two trials on each of three successive days.

(*c*) Step C: Acquisition of the primary avoidance response. During this stage, *S*s were placed in the automatically controlled situation described in Step A, except that the UCS was turned on 5 sec. after the onset of the CSs, as in Step B. As soon as *S* ran to the

opposite compartment, the CSs and the UCS were terminated. The UCS could be avoided by running before it appeared. If *S* made no response, the CSs and UCS continued for 10 sec. after the onset of the UCS. The next trial began 15 sec. after termination of the CSs. It was thus possible for *S* to make no response, to escape once the UCS had occurred, or to avoid the UCS by running before its onset. The level of shock was again just subtetanizing. Each *S* was placed in this situation for 20 trials, twice each day, in the morning and evening, until it reached the criterion of 10 consecutive avoidance responses within a single session. An average of seven sessions was required for this phase.

(*d*) Control procedures for Stage I. The experimental *Ss* served as their own controls during Step A, in which the CSs appeared without the UCS. In addition, eight other *Ss* comprising Control Group A were subjected to the procedure in Step C, except that the UCS was never turned on.

Stage II: Acquisition of the secondary avoidance response.—After acquiring the primary avoidance response (running), Groups I and II were taught the secondary avoidance response (lever pressing).

(*a*) Acquisition. The *S* was placed in one compartment of the shuttle box which had been fitted with a lever. The primary running response was blocked by a barrier in front of the guillotine door. The CSs were turned on after *S* was placed in the compartment and remained on except when the lever was depressed. The UCS was not used at any time during this phase of the experiment. Each *S* was placed in this situation for a 5-min. period on each of seven consecutive days, as soon as it had met the criterion for learning the primary running response. The measure of learning was the percentage of time *S* kept the CSs off by depressing the lever.

(*b*) Control procedures for Stage II. Three Control groups of eight *Ss* were used to test whether *Ss* that had undergone different antecedent conditions would still learn to keep the lever depressed. Group A, described previously, was placed in the secondary situation once per day for seven consecutive days. These *Ss* had previously been subjected to the CS complex without the UCS. The eight completely naive *Ss* of Group B were placed in the secondary situation 5 min. per day on seven consecutive days.

The *Ss* in Group C were taught the running response exactly as were *Ss* in Groups I and II, except that the CS consisted of the opening of the door alone, without the buzzer or light. After reaching the criterion for the running response, they were placed in the secondary situation. They were treated here exactly as were *Ss* in Groups I and II, except that the CS here consisted only of the buzzer and light; the door did not open behind the barrier.

Stage III: Extinction of the primary avoidance response.—Immediately following the acquisition stages, *Ss* in Groups I and II were divided into four groups, matched by pairs for rate of acquisition of the primary running response, level of performance of the secondary lever-pressing response, and sex. Groups I_A (16 *Ss*) and II_A (12 *Ss*) then underwent experimental extinction of the primary response while Groups I_B (16 *Ss*) and II_B (12 *Ss*) remained in their cages. The extinction procedure involved a large number of massed trials during 1-hr. sessions, once a day, in a situation identical to that used for the acquisition of the primary running response except, of course, that the UCS was absent. Extinction was considered complete when *S* ran during none of the first 10 trials of the extinction sessions on both of two successive days and when he ran fewer than 5% of the trials during these two 1-hr. sessions.

Stage IV: Extinction of the secondary avoidance response.—Group I_A and its companion, Group I_B, were now treated differently from Group II_A and its companion, Group II_B. The lever-pressing responses of Groups I_A and I_B were subjected to an experimental extinction procedure in a situation identical to that employed in the acquisition phase (Stage II) of this response. The *Ss* were placed for 5-min. periods, twice a day, for 13 days, into one compartment in which the door was sealed off and the lever inserted. This lever directly controlled the activation of the CSs (buzzer, light, sound of door). There were 25 sessions, or 125 min. of extinction. Neither Group II_A nor Group II_B was subjected to this extinction procedure.

Stage V: Repeat of primary extinction procedure.—The *Ss* in each of the four experimental groups were subjected to the conditions under which the primary running response had been extinguished, a repeat of the procedure for Groups I_A and II_A but not for I_B and II_B. Again the extinction procedure involved a large number of massed trials during 1-hr. sessions, once a day, in a situation identical to the acquisition phase except that the UCS was absent. The rigorous extinction criterion remained the same.

Stage VI: Repeat of extinction of secondary response.—Groups I_A and I_B were returned to the secondary situation for 5-min. periods, twice per day. A total of seven sessions in the lever-pressing situation were carried out during this stage.

Table 1 presents the experimental design in schematic form for all seven groups.

TABLE 1

ARRANGEMENT OF GROUPS FOR EXPERIMENTAL STAGES

Stage	Groups						
	Experimental				Control		
	I_A	I_B	II_A	II_B	A	B	C
I. Prim. Acquisition	X	X	X	X	X[a]	—	X[b]
II. Sec. Acquisition	X	X	X	X	X	X	X[c]
III. Prim. Extinction$_1$	X	—	X	—	—	—	—
IV. Sec. Extinction$_1$	X	X	—	—	—	—	—
V. Prim. Extinction$_2$	X	X	X	X	—	—	—
VI. Sec. Extinction$_2$	X	X	—	—	—	—	—
Ss	16	16	12	12	8	8	8

[a] Shock omitted.
[b] Light and buzzer omitted.
[c] Movement and sound of door omitted.

RESULTS

The results for Steps B and C of Stage I, the acquisition of the primary response, are presented in Fig. 1. The ordinate represents the mean latencies of the response, and the abscissa represents successive groups of trials. The criterion of acquisition was reached by *Ss* of Group I at a mean trial of 130 (range = 66–216). The mean trial at which the criterion was reached by Group II was 145 (range = 86–206).

Also presented in Fig. 1 are the results for the eight *Ss* in Control Group A. These *Ss* underwent an experimental procedure exactly the same as described in Step C above, except that the UCS was not employed. An inspection of Fig. 1 reveals that learning of the primary response occurred only when the UCS was a part of the experimental situation. None of the *Ss* in Control Group A showed any tendency to learn the primary response; similarly, none of the *Ss* in either of the experi-

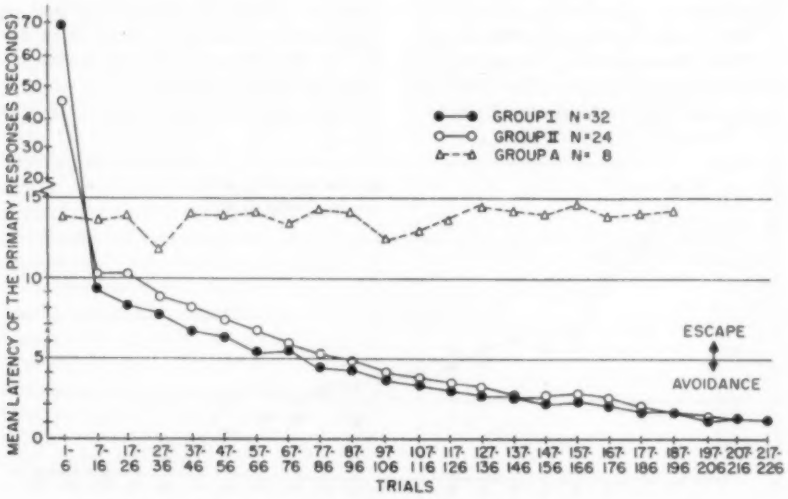

FIG. 1. Stage I, Steps B and C. Acquisition of the primary response for Groups I and II and Control Group A.

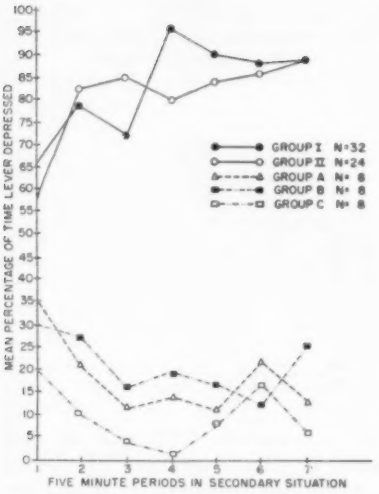
MEAN PERCENTAGE OF TIME LEVER DEPRESSED
FIVE MINUTE PERIODS IN SECONDARY SITUATION
GROUP I N=32
GROUP II N=24
GROUP A N= 8
GROUP B N= 8
GROUP C N= 8

FIG. 2. Stage II. Acquisition of the secondary response for Groups I and II and for Control Groups A, B, and C.

mental groups demonstrated any tendency to learn during Step A, pre-experimental adaptation to the apparatus.

The results for Stage II, the acquisition of the secondary response, are presented in Fig. 2. The ordinate represents the percentage of time *S* spent depressing the lever, while the abscissa represents successive 5-min. sessions.

Also presented in Fig. 2 are the results of the several control procedures employed. Inspection of this figure reveals that lever pressing as a response to the CS-complex was quickly learned and stabilized at a high level for those *S*s who had learned to run from one compartment to the other as a response to the same CS-complex. There was no demonstrable tendency to learn the lever-pressing response without previous experience with the combination of the CS and the UCS. The *S*s in Control Group A, which had had no previous experience with the UCS showed no tendency to learn the lever-pressing response. Similarly *S*s in Control Group C which had had no previous experience with the buzzer and light (but which had learned to run in response to the opening of the door) showed no tendency to learn. Those *S*s in Control Group B, which had had no experience with either the CS or the UCS did not, of course, learn the lever-pressing response.

The findings for Stage III are

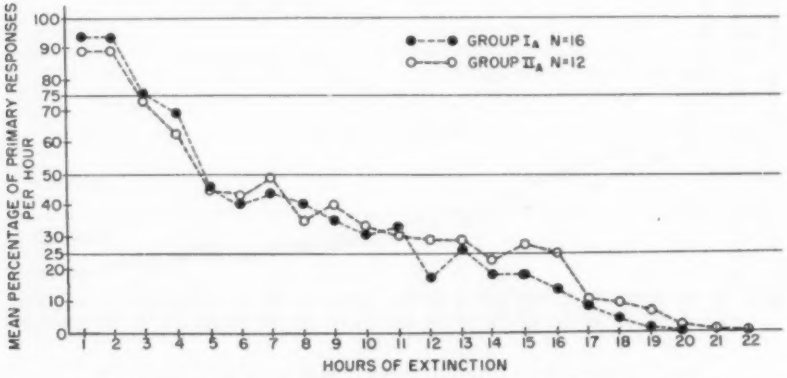

FIG. 3. Stage III. Extinction of the primary response for Groups I_A and II_A.

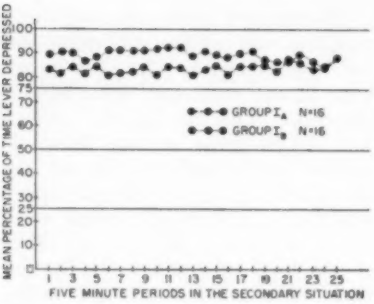

FIG. 4. Stage III. Extinction of the secondary response for Groups I_A and I_B.

presented in Fig. 3. The plotted values represent the mean percentage of trials during which the animals performed the primary running response within 15 sec. of the onset of the CS, by successive hours of extinction. The mean number of trials required to meet the criterion for Group I_A was 2,101 (range = 647–3,812). For Group II_A, the mean was 2,680 (range = 679–4,122). The difference between groups was not statistically significant.

Inspection of Fig. 3 reveals that, in contrast with the fairly rapid acquisition of the primary running response, its extinction involved a very large number of trials spread over a protracted period.

The results for Stage IV, the extinction of the secondary response, are presented in Fig. 4. The ordinate represents the mean percentage of time during which *S*s depressed the lever, while the abscissa represents consecutive 5-min. sessions. Figure 4 presents data for Group I_A which had previously acquired the primary and secondary responses and had extinguished the primary response, and data for Group I_B which had acquired the primary and secondary responses but *had not* extinguished the primary response.

There was no marked difference between the performance of Groups I_A and I_B in Stage IV. The average of the percentages over the 25 sessions was calculated for each *S* in both groups. The mean of these averages in Group I_A was 83.4% (range = 54%–97%; the mean for Group I_B was 89.7% (range = 52%–98%). The difference between these two groups was not statistically significant. It was thus clearly demonstrated that rigorous extinction of the primary running response did not markedly reduce the rate at which animals performed the secondary lever-pressing response.

The findings from Stage V, the repeat of the extinction of the primary response, are presented in Fig. 5. The plotted values again represent the mean percentage of trials during which *S*s performed the primary running response within 15 sec. of the onset of the CS by successive hours of extinction.

The statistical tests pertaining to the results graphically presented in Fig. 5 are presented in Table 2. Inspection of data presented in Fig. 5 and Table 2, together with that presented in Fig. 3, shows that Groups I_A and II_A recovered the response to a

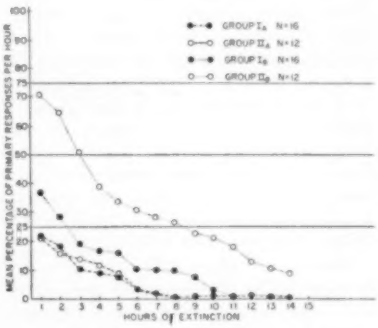

FIG. 5. Stage V. Repeat of the primary extinction procedure for all groups.

TABLE 2

ANALYSIS OF DIFFERENCES BETWEEN MEANS OF GROUPS I_A, I_B, II_A, AND II_B ON THE REPEAT OF PRIMARY EXTINCTION PROCEDURE

Difference	*df*	*t*
$I_A - I_B$	15	1.36
$I_A - II_A$	26	.07
$I_A - II_B$	26	4.35***
$I_B - II_A$	26	.06
$I_B - II_B$	26	2.56*
$II_A - II_B$	11	4.33**

* $P < .02$.
** $P < .01$.
*** $P < .001$.

substantial degree in spite of the fact that this response had previously been extinguished to the free-operant level. The *Ss* in Group II_A, who rested in their cages during the extinction procedure for the secondary response, exhibited the primary response just as frequently as did *Ss* in Group I_A, who had participated in the secondary extinction procedure.

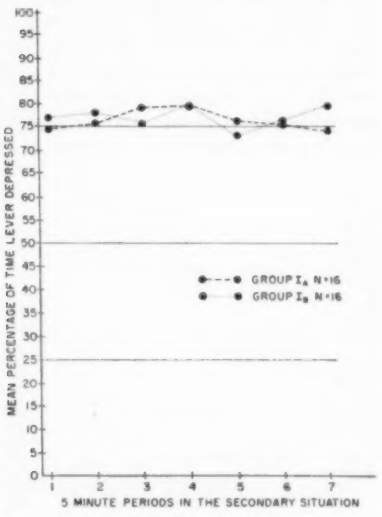

FIG. 6. Stage VI. Repeat of the secondary extinction procedure for Groups I_A and I_B.

The strength of the primary running response was considerably diminished for *Ss* in Group I_B after their experience in the secondary lever-pressing situation. This diminution, however, was not as great as that shown by groups which had experienced previous primary extinction (Groups I_A and II_A). The three groups which had undergone extinction procedures of one kind or another (I_A, I_B, and II_A) all showed a marked reduction in the tendency to respond in the primary situation when compared with the one group which had experienced no extinction procedure (II_B).

TABLE 3

ANALYSIS OF DIFFERENCES BETWEEN MEAN PERFORMANCE OF THE SECONDARY RESPONSE IN STAGES II, IV, AND VI, FOR GROUPS I_A AND I_B

Group	Mean	Mean	Diff.	*df*	*t*
I_A	90.0% (II)	83.4% (IV)	6.6%	15	.87
I_B	89.3% (II)	89.7% (IV)	.4%	15	.19
I_A	83.4% (IV)	76.3% (VI)	7.1%	15	3.52***
I_B	89.7% (IV)	77.2% (VI)	12.5%	15	12.90***

*** $P < .001$.

The findings for Stage VI, the repeat of the extinction of the secondary response, are presented in Fig. 6. Group I_A, which was subjected to all of the acquisition and extinction procedures, is compared in this graph with Group I_B, which had experienced all of the stages of the experiment except Stage III, extinction of the primary response.

Statistical analyses of differences in the performances in the three secondary situations (Stages II, IV, and VI) are presented in Table 3. These were calculated on the basis of the final three sessions of the acquisition phase (Stage II) and the whole of the extinction phases (Stages IV and VI).

It is apparent, from an inspection of Fig. 2, 4, and 6, and from Table 3, that while there was no marked decrease in the strength of the secondary response following the initial primary extinction procedure for Group I_A, both Groups I_A and I_B did show a significant diminution in the strength of the secondary response following the repeat of the primary extinction procedure.

Thus, while marked extinction of the secondary lever-pressing response was not accomplished during the time given to this procedure, the evidence indicated that this response was not impervious to extinction effects.

Discussion

The results of the present experiment provide data which are pertinent to the empirical question of whether a response may persist without apparent motivation. An analysis of the results of the various attempts at experimental extinction indicated that the primary running response was susceptible to extinction effects and that the secondary lever-pressing response was very resistant to extinction. All *Ss* subjected to experimental extinction eventually ceased running in response to the CSs. It is true that this primary response was also extremely resistant to extinction, the number of trials required to reach the criterion ranging from 647 to 4,122, with a mean of 2,541. It is also true that the running response recovered to some extent in some *Ss* after it was thought to be completely extinguished. The probability of *S*'s running from one compartment to the other as a response to the CSs did, however, decrease significantly, and there is no reason to suppose that the probability of this response could not have been brought to and stabilized at the free-operant level.

The *Ss* who were subjected to the primary extinction procedures persisted in the secondary lever-pressing response to an extent equal to that of *Ss* who had no primary extinction experience at all. It is clear, then, that a secondary response may persist after the CS which had the power to cue and motivate both this secondary response and a primary response, has for the most part lost its power to elicit the primary response. If one assumes that extinction of the running response was satisfactory evidence that the acquired fear drive had been correspondingly reduced, then this finding constitutes an experimental demonstration of the persistence of an instrumental response without apparent motivation. The result resembles the phenomenon which has been referred to as "functional autonomy." That is, a response (lever pressing) persisted over a protracted period even though the basis for this response (the acquired fear drive) had apparently been removed.

This experiment did not determine whether or not the secondary response was susceptible to total extinction. While the amount of the lever-pressing response was not markedly reduced in either Stage IV or Stage VI, this does not mean, of course, that the strength of this response could not eventually have been brought to the free-operant level. The amount of lever-pressing activity did decrease significantly from Stage II to Stage VI, although there is no indication of progressive decrease in response strength *within* either Stage II or Stage IV. We can only conjecture about what might have happened within longer protracted extinction trials; extinction might eventually have set in, or the response level might have fallen with repeated extinction of the primary response.

The results of the present experiment can be explained in several different ways. First, one might hypothesize that the extinction of the primary response did not result in a reduction in the strength of the fear drive motivating the secondary response because the two situations were easily discriminated. The generalization of the positive avoidance response to fear may have been greater than the generalization of extinction. Second, it might be argued that

the secondary response became functionally independent of the fear drive responsible for its acquisition because the lever pressing became in some way a self-rewarding activity. Third, it could be maintained that the drive concept is perhaps not necessary to explain the acquisition, persistance, or extinction of a response.

To the present writer, these results do not seem to necessitate the complete abandonment of a drive-reduction orientation. It is possible that the extinction of the primary running response represented a reduction in the strength of the fear drive to a point at which this drive, together with the habit to run, was no longer capable of eliciting the running response. Enough fear remained, however, so that this drive could, in conjunction with the habit to press the lever, elicit the lever-pressing response. Whatever the theoretical explanation, it seems clear that the principal aim of the investigation has been achieved, namely the experimental demonstration of the persistence of a response without apparent motivation, under conditions that do not permit easy explanation according to the ad hoc principles usually invoked.

Summary

The persistence of a lever-pressing response in the absence of apparent motivation was demonstrated. The *Ss* were 80 rats, in four experimental and three control groups. The *Ss* first learned a primary response, running to the opposite compartment of an automatic shuttle box, to avoid the UCS (just subtetanizing shock), and then a secondary response, lever pressing, to avoid the CSs (light, buzzer) with shock absent. After rigorous extinction of the running response, the lever pressing continued unabated. Some spontaneous recovery of the running response occurred, but lever pressing continued with substantial strength after running was again extinguished.

References

Allport, G. W. *Personality.* New York: Holt, 1937.

Hull, C. L. *A behavior system.* New Haven: Yale Univer. Press, 1952.

McClelland, D. C. Functional autonomy of motives as an extinction phenomenon. *Psychol. Rev.*, 1942, **49**, 272–283.

Miller, N. E. Studies of fear as an acquirable drive: I. Fear as motivation and fear-reduction as reinforcement in the learning of new responses. *J. exp. Psychol.*, 1948, **38**, 89–101.

Robinson, H. B. The persistence of a response in the absence of apparent motivation. Unpublished doctoral dissertation, Stanford University, 1957.

Woodworth, R. S. *Dynamic psychology.* New York: Columbia Univer. Press, 1918.

(Received April 9, 1960)

Journal of Experimental Psychology
1961, Vol. 61, No. 6, 489–493

HUMAN HEART RATE RESPONSES DURING EXPERIMENTALLY INDUCED ANXIETY[1]

GEORGE E. DEANE[2]

University of Connecticut

One proposed paradigm for the experimental establishment of anxiety is the pairing of an initially neutral stimulus with an aversive stimulus. Electric shock has been employed commonly as the aversive stimulus and various autonomic responses have been used as the behavioral manifestation of anxiety. When changes in heart rate have been employed, some studies have reported cardiac acceleration while others have reported deceleration in the interval between warning signal and shock. Typical of the deceleration findings is the report of Notterman, Schoenfeld, and Bersh (1952) who used trace conditioning and found a cardiac slowdown during the 6-sec. interval between warning signal and shock. Representative of the opposite finding is the report of Lacey and Smith (1954) of cardiac acceleration following verbal stimuli which had been associated with shock.

Zeaman and Wegner (1957) reported finding both effects in different *Ss* and suggested that perhaps differential breathing patterns might be related to the dual cardiac effects. In a test of this hypothesis, Deane and Zeaman (1958) warned *Ss* that they would receive a shock at the end of a series of visually presented numbers. It was found that looking at the numbers for the first time brought about an acceleration in rate, but that after one shock the acceleration disappeared and was replaced by deceleration late in the series of numbers. None of the respiration measures employed correlated with the observed cardiac changes. The authors tentatively concluded that the shock itself was responsible for the loss of acceleration and the acquisition of the deceleration near the point where shock had been received.

The present study is an extension of the study by Deane and Zeaman (1958) and is designed to discover the controlling conditions of these two cardiac effects. Specifically, it attempts to answer the following questions: (*a*) Is the acceleration reported in the previous study merely a transitory effect which would disappear soon after the initial warning of forthcoming shock, or is it necessary for *S* actually to receive a shock before it disappears? (*b*) Does the deceleration which occurred just prior to and during the time shock was expected depend upon actually having received a shock, or are verbal warnings sufficient to produce it?

METHOD

Apparatus.—The apparatus consisted of an electrocardiograph and a 23-VAC shock source, both of which were described in an earlier paper (Zeaman, Deane, & Wegner, 1954). A memory drum was used to present the numbers 1 through 12 at the rate of one every 3 sec. All apparatus with the exception of the shock electrodes, cardiograph leads, and memory drum was outside *S*'s cubicle.

[1] This paper is based on a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the University of Connecticut. The writer wishes to express his gratitude to David Zeaman for his advice throughout the execution and writing of this study.

[2] Now at Harpur College.

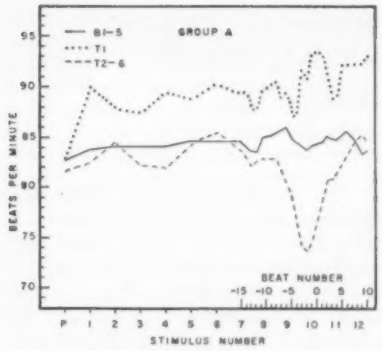

FIG. 1. Mean heart rate during the number stimuli with trials as the parameter: Group A. (The *x* axis is a time line calibrated in terms of both stimulus number and ordinal beat number—see text.)

Subjects.—The *S*s were 36 male students at the University of Connecticut. All were paid volunteers and were unaware that shock was to be involved in the experiment.

Base-level measurements.—Each *S* was seated in the cubicle facing the memory drum and, while the cardiograph leads were being placed on him, was told that the purpose of the experiment was to study physiological changes that may occur when people attend to a series of visual stimuli. He was told to watch the window of the memory drum where he would see the numbers "1" through "12" appear in order several times. The *E* then left *S*'s cubicle and after about 3 min. began the first of five presentations of the numbers for base-level measurements of cardiac activity. The series were spaced at irregular intervals of from 1½ to 2½ min.

Anxiety measurements.—At the end of the fifth base-level presentation *E* entered *S*'s cubicle and, depending upon the group to which *S* had been assigned randomly, gave him one of four sets of instructions.

In Group A each *S* was told he would continue to see the same series of numbers, but that now when "10" appeared he would receive a mild electric shock on the first two fingers of his left hand where *E* was placing two electrodes. It was emphasized that *S* might not receive shock every time "10" appeared, but that he would never receive a shock unless "10" was in the window. Actually, only the first of the six following presentations involved a 3-sec. shock during "10."

Group B received exactly the same instructions as Group A. This group, however, never received a shock on any of the six presentations.

In Group C each *S* was told he would continue to see the same series of numbers but that now when *some* number appeared in the window he would receive a mild electric shock on the first two fingers of his left hand where *E* was placing two electrodes. It was emphasized that *S* might not receive shock on every series of numbers, but that he would never receive shock unless *some* number was in the window. Actually, only the first of the following six presentations involved a 3-sec. shock during number "10".

Group D received exactly the same instructions as Group C. This group, however, never received a shock on any of the six presentations.

Heart rate was recorded on each trial from about 2 min. before "1" appeared to about 2 min. after "12" had passed.

RESULTS

Plotting the form of the response.—In Fig. 1 through 4 are plotted the results for each of the four groups. In every case the solid line is a mean of the five base-level presentations (B1–5), the dotted line is the first shock anticipation trial (T1), and the dashed line is a mean of the following five shock anticipation trials (T2–6). For each group the data of Trials

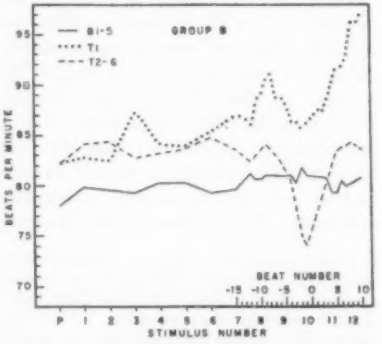

FIG. 2. Mean heart rate during the number stimuli with trials as the parameter: Group B.

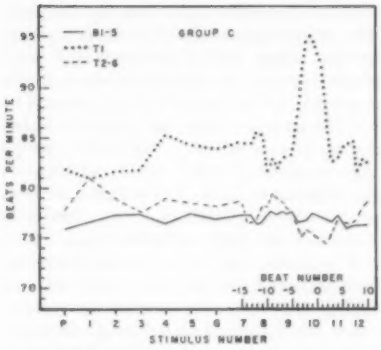

FIG. 3. Mean heart rate during the number stimuli with trials as the parameter: Group C.

B1–5 were pooled for a measure of base level and the data of Trials T2–6 were pooled because the differences among the trials were small and unsystematic. The *y* axis is heart rate in beats per minute. At point P on the *x* axis is plotted the mean of the three beats immediately preceding the appearance of Stimulus "1," and at each of the points labeled Stimuli "1" through "6" is plotted the mean of the middle three beats occurring during presentation of these numbers.

In order to preserve the exact form of the response in the vicinity of the point where shock was to be expected, a beat-by-beat analysis of rate was made for the 15 beats preceding the end of Stimulus "10," and the 10 beats following the end of Stimulus "10." These are labeled beat numbers −15 through −1 and 1 through 10, respectively. Where this beat-by-beat analysis was made, the average stimulus number present during these beats is noted.

Analysis of combined groups.—In order to describe what happened in anticipation of shock early in the number series independently of the changes that occurred in the immediate vicinity of and during the shock itself, one analysis was made for the first eight stimulus numbers and one for the numbers "9" through "12."

Figures 1 through 4 indicate that for all groups the initial effect of instructions regarding impending shock was to accelerate the rate over the base level. An analysis of variance was performed upon the mean rate during Stimuli "1" through "8" for Trials B1–5 and Trial T1 and, as shown in Table 1, the Trials effect is significant at better than the .01 level.

Figures 1 through 4 also show that this acceleration was not a transitory effect which disappeared soon after the initial warning of forthcoming shock. Only in Groups A and C which received shock was there any indication of the rate dropping back to the base level during Trials T2–6. This is supported by the results of an analysis of variance performed upon the mean rate during Stimuli "1" through "8" for Trials B1–5 and T2–6. Table 1 shows that both the Trials effect and the Trials × Shock interaction are significant at better than the .05 level.

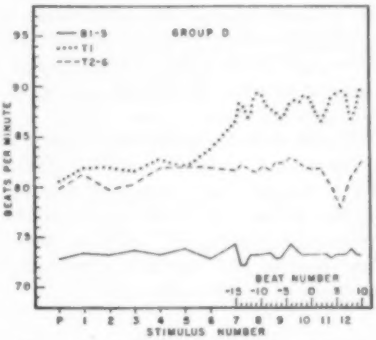

FIG. 4. Mean heart rate during the number stimuli with trials as the parameter: Group D.

TABLE 1

ANALYSIS OF VARIANCE OF HEART RATE LEVEL DURING STIMULI "1" THROUGH "8"

Source	*df*	Trials B1–5 and T1		Trials B1–5 and T2–6	
		MS	*F*	*MS*	*F*
Between					
Instructions (I)	1	473.30	1.81	495.62	2.65
Shock (S)	1	145.01		25.28	
I × S	1	23.88		14.07	
Error	32	262.11		187.03	
Within					
Trials (T)	1	755.52	8.91**	158.12	5.57*
T × I	1	38.38		46.91	1.65
T × S	1	9.91		145.36	5.12*
T × I × S	1	25.78		4.37	
Error	32	84.80		28.39	

* $P < .05$.
** $P < .01$.

The results in Fig. 2 for Group B during Trials T2–6 show quite clearly that a deceleration can occur just prior to and during the time shock is expected even though shock has never been received. In order to evaluate the effects of instructions and shock upon the observed deceleration, a 2 × 2 analysis of variance was performed on the mean difference scores (mean rate during Stimuli "7" and "8" minus mean rate during Stimuli "9" and "10") during Trials T2–6. The Instructions effect is significant at better than the .01 level ($F = 37.8$; $df = 1/32$) and Shock is significant at better than the .05 level ($F = 5.4$; $df = 1/32$). The Instructions × Shock interaction gives an $F < 1$.

Figure 4 reveals that for Group D there was a drop in rate on Trials T2–6 during "12," the last number in the series. A t test between the mean rate during "8" and the mean rate during "12" gives a two-tailed $P < .02$ ($t = 2.98$; $df = 8$).

DISCUSSION

The present study indicates that there are two opposing heart rate responses during experimentally induced anxiety. The acceleration which occurred in all groups after shock instructions were given was not transitory but was a stable effect which remained during Trials T2–6. The amount of acceleration remaining during these trials was less for the groups that received shock. An experiment is in progress to test the hypothesis that the acceleration will decrease only if *S* receives a subjectively mild shock.

The results also indicate that deceleration occurred immediately prior to and during the time *S* expected to receive shock even though shock had not been received. Lacey and Smith (1954) who reported only acceleration in the interval before shock would not have discovered this deceleration in their data if it were present since they used as a response measure the mean of the six fastest beats in the interval. Group D which was told it would receive shock on some number, although actually never receiving one, exhibited a significant deceleration with respect to the immediately preceding rate during "12" on Trials T2–6. Although this deceleration had not been predicted, it seems possible that the passing of each successive number without shock increased the probability of *S*s expecting it later in the series, and after 11 stimulus numbers passed without shock they were forced to the conclusion that shock must be coming during "12", the last number in the series.

It is interesting to note that verbal instructions regarding when to expect shock were apparently more effective in bringing about the deceleration than actually receiving the shock itself. Group B was told to expect shock during "10" and in spite of the fact it was never received, showed a mean deceleration of 4.4 beats per minute during Trials T2–6. Group C, although not told when to expect shock, actually received one during "10" and yet exhibited a mean deceleration of only 1.2 beats per minute. A t test between the amount of deceleration in these groups is significant at better than the .05 level ($t = 2.33$; $df = 16$).

The present results show that both

acceleration and deceleration can occur *without* shock ever being administered. This would indicate that these cardiac effects may be transferred or generalized responses dependent upon *Ss'* previous conditionings with respect to shock and verbal warnings of shock. If they are transferred responses, it is necessary to demonstrate how they were initially acquired, and the difficulties of the stimulus substitution and Hullian reinforcement positions for explaining at least the deceleration have been noted in a previous paper (Zeaman et al., 1954). It might be speculated instead that these two cardiac effects are perhaps unlearned responses associated with what may be called anxiety and fear with the following empirical significance. When *S* expects a noxious stimulus of unknown strength, a state of anxiety with its associated response of cardiac acceleration is aroused a relatively long time before the stimulus is expected. In addition, if *S* expects the noxious stimulus at a particular instant in time a state of fear with its associated response of cardiac deceleration is aroused immediately prior to and during the time the stimulus is expected. If the stimulus is found to be subjectively mild, the anxiety is reduced, but it is apparently the case that even experiencing the noxious stimulus as relatively mild is not sufficient to abolish the fear response at the expected locus of the stimulus. The generality of these speculated distinctions between anxiety and fear is at present being tested using noxious stimuli other than shock.

Summary

The present study was designed to discover some of the controlling conditions of the two cardiac effects found during anxiety induced by signals of shock-to-come. Varying instructions were given to four groups of male students. In all groups there was an acceleration in cardiac rate on the first shock anticipation trial. This acceleration was not a transitory effect which disappeared soon after the initial warning of forthcoming shock, although the two groups that received shock did not retain as much of the acceleration as did the two groups not receiving shock. The groups that were told when to expect the shock showed in addition a pronounced deceleration at the time shock was expected. Both acceleration and deceleration occurred in the group that never received a shock. It was speculated that these two cardiac effects may be unlearned responses associated with what may be called anxiety and fear. When *S* expects a noxious stimulus of unknown strength a state of anxiety with its associated response of cardiac acceleration is aroused, and in addition if *S* expects the noxious stimulus at a particular instant in time a state of fear with its associated response of cardiac deceleration is aroused immediately prior to and during the time the stimulus is expected.

References

Deane, G. E., & Zeaman, D. Human heart rate during anxiety. *Percept. mot. Skills.*, 1958, **8**, 103–106.

Lacey, J. I., & Smith, R. L. Conditioning and generalization of unconscious anxiety. *Science*, 1954, **120**, 1045–1052.

Notterman, J. M., Schoenfeld, W. N., & Bersh, P. J. Conditioned heart rate response in human beings during experimental anxiety. *J. comp. physiol. Psychol.*, 1952, **45**, 1–8.

Zeaman, D., Deane, G. E., & Wegner, N. Amplitude and latency characteristics of the conditioned heart response. *J. Psychol.*, 1954, **38**, 235–250.

Zeaman, D., & Wegner, N. A further test of the role of drive reduction in human cardiac conditioning. *J. Psychol.*, 1957, **43**, 125–133.

(Received April 14, 1960)

Journal of Experimental Psychology
1961, Vol. 61, No. 6, 494–500

EFFECTS OF DENSITY ON IDENTIFICATION AND DISCRIMINATION IN VISUAL SYMBOL PERCEPTION [1]

WARREN H. TEICHNER, RAYMOND REILLY, AND ERNEST SADLER

University of Massachusetts

Among the experimental procedures used in traditional studies of attention span and more recent studies of human information transmission, are experiments which require *S* to identify the different, discrete stimulus categories or symbols (e.g., letters, shapes, numbers) present on a briefly exposed display. In such studies the symbols are arranged on the display in a manner such as to minimize any need for *S* to search the display in order to detect the symbols. Such studies might be called identification studies as distinct from search studies which have varied the spatial location of a single critical symbol (often hidden among a set of irrelevant symbols) and measured *S*'s search time. No research seems to have been done with tasks which combine both requirements, that is with tasks which require *S* to search and then identify more than one symbol on a briefly exposed display. Yet such tasks are characteristic of many actual situations including air traffic control and a variety of multisignal operating conditions. It was our purpose, therefore, to investigate performance in this type of task.

Although no direct information about performance in this type of task appears to be available, it seems reasonable to generalize as hypotheses the principles which have been developed to predict performance in the more simple identification tasks. In particular it has been found (e.g., Anderson & Fitts, 1958; Miller, 1956; Shannon & Weaver, 1949) in identification that some form of redundancy often increases the amount of information that can be transmitted (number of categories that can be identified). It has also been found that increasing the number of stimulus categories (message length) first increases and then decreases information transmission (Anderson & Fitts, 1958; Miller, 1956). On the other hand, increasing the complexity of the display in this manner produces only a decrease in performance when the dependent variable is the number of correctly identified symbols (Anderson & Fitts, 1958; Woodworth & Schlosberg, 1954). Generalizing these principles to the present task it might be expected that redundancy in some form would aid performance and complexity would act in general as a decremental factor, with the qualification noted.

One way of introducing redundancy in a search-identification task is simply to repeat the symbols at different parts of the display. During a brief exposure this should increase the probability of detection and thereby increase the number of categories that can be identified. It would seem from this that for a given number of categories, identification should be positively related to the frequency or density of repetition. On the other hand, as the total number of symbols (Density × Categories) on the dis-

[1] This research was carried out under Contract N61339-588 with the United States Naval Training Device Center. Presented to the Eastern Psychological Association, April 1960.

play increased, performance should decrease as a complexity effect. Thus, for any given number of categories there should be an optimum density and the optimum density should be inversely related to the number of categories on the display to be identified. This hypothesis was one major concern of this study.

The *S*'s identification of categories describes only one kind of information being processed. Either explicitly or implicitly, he is also processing information concerning the *number* of categories present. Even if his identifications were completely wrong, they would still represent an estimate of the number of categories he has discriminated as different and present on the display. Miller (1956) makes this distinction in terms of the span for perception and the span for memory. Since the perceptual process is basic to the identification, the maximum amount of information that *S* can handle as an ideal is his perception span or perceptual capacity; any loss in transmitted information must be the result of the requirement for identification. Identification, however, requires only that *S* recall what he has perceived. Thus, when the amount of information transmitted is less than *S*'s perceptual capacity it must be the result of a limit in his short-term memory capacity. This assumption is reasonable if it can be shown that the number of categories that *S* can count is greater than the number he can name. This hypothesis was the second major concern of the present study.

There is some indirect evidence which supports our second hypothesis. When studies of attention span for the naming of different symbols are compared with the attention span for numerosity, it is found that the latter span is greater (Woodworth & Schlosberg, 1954). However, numerosity studies typically vary the density of a single category, e.g., number of dots, and require *S* to judge the number whereas naming studies vary the number of categories, but have not presented any category on the display more than once. Thus, available comparisons of number estimates and naming estimates do not provide a direct test of our hypothesis. Such a test requires number and naming judgments of the same displayed material. The present study provided for this under the conditions demanded by our first hypothesis. In addition, for both types of information processing, we were also concerned with the possible sequence effects which might arise if *S* were presented with a series of displays whose densities were random or grouped in the series.

Method

Apparatus.—Two hundred negative slides were made by photographing capital latin letters which were typed on 3 × 5 in. cards. All letters of the alphabet were used except the letter Q. The slides varied along two dimensions: (*a*) with each particular letter representing a category, the slides ranged from three to six categories; (*b*) with density defined as the amount of replication or frequency of a particular letter, the slides ranged in density from one to five. Density per category was constant within any individual slide. Figure 1 presents two examples of slides actually used. The first display has four categories and a density of three. The lower display has three categories and a density of four. A display with only one of each category has a density of one.

In all there were a total of 20 density-category combinations (5 different Densities × 4 different sets of Categories). Each of these combinations was replicated 10 times with a different random sample of the alphabet to make up the 200 slides.

To assign the location of the letters on the slide, the central area of the 3 × 5 in. card photographed was divided into a 10 × 10 matrix each cell of which was the space

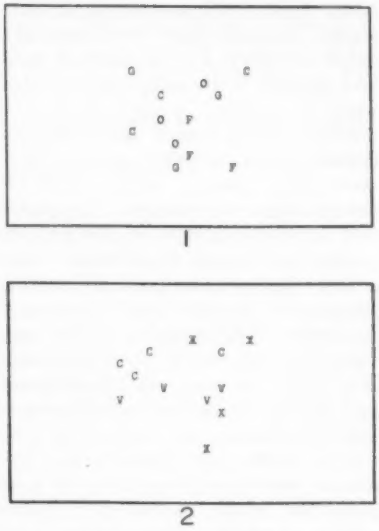

FIG. 1. Examples of two displays used. (1—Categories = 4, Density = 3; 2—Categories = 3, Density = 4.)

required by the typewriter to print a letter. The cells of the matrix were so selected that each one was separated from adjacent ones by one typewriter space both horizontally and vertically. The selection and assignment of letters to the 100-celled matrix so formed was then a two-step process: (*a*) random selection of different letters from the alphabet; (*b*) random assignment of letters to the matrix cells.

The exposure time of each slide was 1 sec.; the interslide interval was 10 sec. *Ss* sat in student arm-type chairs in groups of three such that the average distance from their eyes to the projected letters was 82 in., the average illumination at *S*'s position was 0.3 ft-c. The projected letter size was 0.5 in.; the total projected area on the screen was 15 × 25 in. and the illumination of this area was 0.21 ft-c.

Procedure.—The first session was 30 min. and was devoted to instructions and practice with a sample of slides not used later. On three successive days following this, *Ss* were presented the 200 slides in sessions of approximately 1 hr. Each of these sessions presented the slides in a different sequence. One sequence was a completely random arrangement of all slides (random presentation). In another sequence all slides of Density 1 were presented first, those of Density 2 second, etc. (systematic presentation). In this sequence, categories were randomized within each density group. A third sequence was the same as the second except that the order of presentation of the density groups was randomized (partly systematic presentation). The order in which *Ss* experienced the three sequences was arranged in a latin square. The *Ss* were never informed about the range of densities or categories, nor were they given knowledge of results. They were informed that the letter Q would never be displayed.

One group of 12 *Ss* identified the different letters displayed on each slide without regard to density by writing their judgments on prepared data sheets during the 10-sec. interslide interval. This constituted the identification task group. A second set of 12 *Ss* recording in the same way reported the number of different letters displayed without regard to density. This constituted the discrimination task group.

Subjects.—The *Ss* were volunteer, undergraduate, summer university students with no previous relevant experience. Each group of 12 *Ss* competed among its own members for cash prizes of $20, $10, and $5. The payoff score was calculated as the total percentage correct over the entire experiment.

RESULTS

Identification analysis.—The data were analyzed in terms of three measures: (*a*) the percentage of displayed categories which were correctly identified, (*b*) the percentage number of categories in *S*'s report less than the number displayed (error of omission), (*c*) the percentage number of categories in the report greater than the number displayed (error of commis-

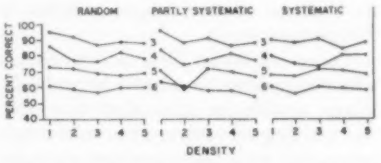

FIG. 2. Effects of density on correct identification for three sequences of presentation. (The parameter is number of categories.)

sion). The results for the percentage of correctly identified categories based on the combined data of all 12 *S*s are shown in Fig. 2. Inspection of this figure reveals that neither methods of presentation nor density had any but random or negligible effects on *S*'s ability to identify the categories. On the other hand, number of categories had a marked, systematic effect. These results appear to be so clear that although the experiment was designed for a specific statistical analysis (analysis of variance of the within-*S*s type), the analysis was not done. The significant variable, categories, was to be expected and even if the other two variables were significant, it is clear that their effects were negligible.

The effect of categories is shown more clearly in Fig. 3 which presents the same data averaged over all densities within each method of presentation. This figure also shows the same treatment of the data for errors of omission and commission. The curves drawn through the points were fitted by inspection. This figure shows clearly that the percentage of correct identification was inversely proportional to the number of categories, that the percentage of omissions was directly proportional to the number of categories and that the percentage of commissions was a negatively accelerated, decreasing function which reached zero by six categories and was almost zero by five. It emphasizes again the failure of method of presentation to affect the results.

Inspection of the two straight lines in Fig. 3 reveals that they have approximately the same slope though different in sign. This is not a necessary consequence of the scoring technique since, for example, the percentage of correct identifications could be the same regardless of the occurrence

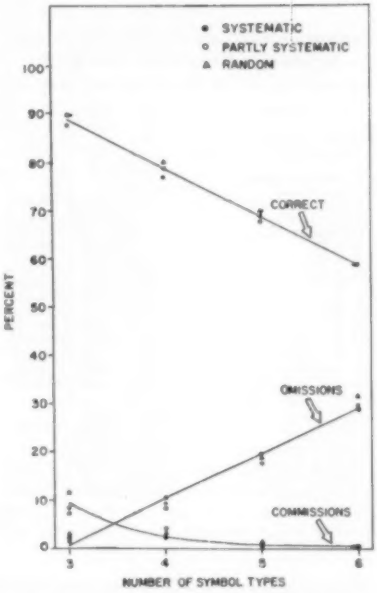

FIG. 3. The effect of categories (number of symbol types) on three aspects of the identification response.

of either of the two types of error. That is, since percentage correct depended on both correctness of identification and number of identifications (up to the number displayed) it was only partly dependent on omissions.

Discrimination analysis.—The dependent measure in this analysis was the number of different letters (categories) reported by *S* rather than the names of these letters. For *S*s performing the discrimination task this was the numerical value they recorded. For *S*s performing the identification task this was the total number of letters recorded for a given display regardless of correctness, omissions or commissions. Table 1 presents a summary of an analysis of variance performed on these measures.

TABLE 1

SUMMARY ANALYSIS OF VARIANCE OF THE NUMBER OF CATEGORIES REPORTED

Source	*df*	*MS*	*F*
Between *S* total (*S*)	23		7.87*
Task (number vs. letter) (T)	1	102.45	
S/T (between error)	22	13.01	
Within *S* total	1416		
Presentation (P)	2	0.825	0.42
Density (D)	4	0.250	0.27
Categories (C)	3	104.650	189.58**
P × D	8	0.367	15.96**
P × C	6	0.280	2.48*
D × C	12	2.140	3.28**
P × D × C	24	0.099	6.60**
P × T	2	0.990	0.50
C × T	3	5.753	10.42**
D × T	4	0.690	0.74
P × C × T	6	0.118	1.04
P × D × T	8	0.400	17.39**
C × D × T	24	0.092	0.14
P × C × D × T	24	.230	15.33**
Within *S* error total			
(*S* × P)/T	44	1.986	
(*S* × C)/T	66	0.552	
(*S* × D)/T	88	0.929	
(*S* × P × C)/T	132	0.113	
(*S* × P × D)/T	176	0.023	
(*S* × C × D)/T	264	0.652	
(*S* × C × D × P)/T	516	0.015	

* $P < .05$.
** $P < .001$.

Inspection of Table 1 indicates that neither method of presentation nor density was a significant main effect, a finding which confirms our earlier judgment regarding these factors in relation to identification. This table also shows that the main effects of task and of number of displayed categories were both significant as was the interaction of these two factors. These results are made explicit in Fig. 4 which presents the mean number of different letters reported by each group as a function of the number of letters displayed. These data are averaged over both density and presentation method. The straight line in this figure represents perfect agreement between the number reported and displayed, or in other terms, perfect information transmission. The differences between this line and the empirical values show the amount and direction of *S*-error, or in informational terms, the equivocation. It may be seen that as the number of categories displayed increased, the mean number reported by both groups changed from an overestimate to an underestimate. For *S*s performing the discrimination task, i.e., reporting number, the greatest accuracy was with displays of approximately 4.2 categories. This can be considered an optimum since the size of the error increases on either side of it. A similar optimum is apparent for the identification task group at approximately 3.4 letters.

Figure 4 also shows that the size of the error, except for the smallest number of categories, was greater when *S*s were required to name the letters. Also, although the equivocation increased for both groups, the increase was somewhat more accelerated for the identification group. The greater flattening of the curve at the high end for this group suggests an approach to a nearer limit.

Graphical analysis shed some light on other significant interactions. Inspection of a plot of presentation method by density indicated that for *S*s reporting numbers, the mean

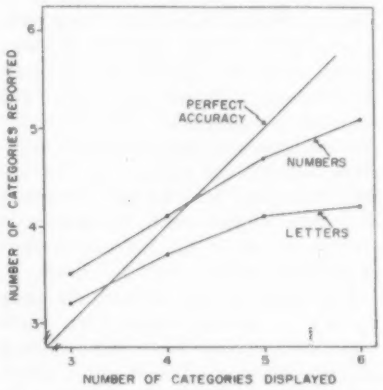

FIG. 4. Number of categories reported as a function of the number displayed for the discrimination task group (numbers) and the identification task group (letters).

error at different densities was slightly greater in order of completely systematic, and random presentations. No consistency was suggested for *Ss* reporting letters.

With regard to the interaction of Presentation × Categories, regardless of density, the amount of error at a given number of categories varied with presentation method. For number-writers, the least error made with three and four categories was in order of completely systematic, partly systematic and random presentations; with five and six categories, the least error was made in order of partly systematic, random, and completely systematic presentations. Again, letter-writers showed no discernible consistency.

Regarding the Density × Categories interaction, both groups made the least error at all categories with density equal to one. Densities greater than one had no further apparent effect. Figure 5 shows this result and in addition presents the data in terms of the amount of information transmitted at the various densities, for the three methods of presentation, for each of the two groups. The remaining interactions were difficult to interpret. However they did suggest that the least-error-effect obtained with a density of one depended in a complex way upon the specific combination of number of categories, method of presentation and task.

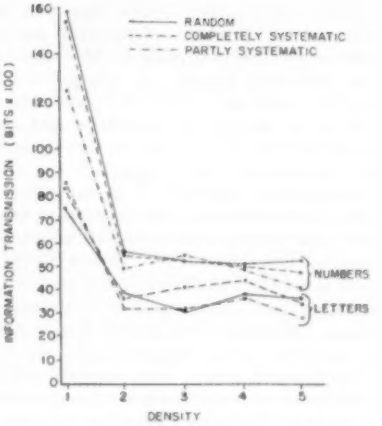

FIG. 5. Information transmitted as a function of density averaged over all categories. (Numbers refers to the discrimination group, letters to the identification group. The parameter is presentation method.)

DISCUSSION

The data are clear in showing that percentage of correct identification is inversely proportional to the number of categories displayed. The identification analysis shows that this increasing loss of information transmission depends not only on increased error in identification but also on an increase in percentage of omissions, i.e., on a concommitant increase in equivocation and decrease in response information.

The results also show that density had no effect on identification and no significant main effect on discrimination. For the latter, a clear difference was found between the effects of a density of one and all other greater densities. These results do not support our hypothesis that redundancy, at least in this form, should aid performance. We suspect that the exposure time may have been a critical factor. Studies have been reported in which density did seem significant (Christner, Schutz, & Ray, 1959; Green & Anderson, 1956). In these experiments density was a decremental influence on performance. However, these were studies in which total density (Density × Categories) was analyzed rather than per category density and in which search time was the dependent variable. It is difficult, therefore, to generalize from these to our type of study.

Our distinction between identification and discrimination revealed itself markedly in terms of differences in the ac-

curacy of number judgments and/or amount of number information transmitted by the two groups, the value of the optimum number of categories for perfect accuracy, and the differential effect of density on numbers and on identified letters, i.e., the former showed an interaction not suggested by the identification analysis. It seems reasonable to conclude on this basis that two different processes are involved although one subsumes the other. Thus a distinction is in order, as Miller (1956) suggests, between perception span and memory span. Although this distinction is not a new one, it has always been only an hypothesized one. The present study is believed to provide a first experimental distinction between the two concepts. In informational terms the results indicate that perceptual capacity (perception span) is greater than short term memory capacity (memory span).

An important implication of this is that *Ss* can be expected to make decisions about displays containing symbols that they can discriminate even if their ability to name the symbols is limited. A related implication is that the absolute number of categories that *S* can discriminate along a single dimension may be limited to approximately seven (Miller, 1956) but is probably importantly greater than this if *S* is not required to name the categories.

Summary

Twenty-four *Ss* viewed a series of 200 slides presented with a 1-sec. exposure. The slides varied in number of different letters (categories) contained and density of the categories. All slides varied randomly in letter location. Half of the *Ss* identified the categories; half of them reported the number of categories displayed. The series of slides was presented in random sequence and two variations of grouped density. For the experimental conditions used, the following conclusions appear warranted:

1. For identification of categories, percentage correct identification is inversely proportional to number of categories; percentage of ommissions is directly proportional to number of categories; percentage of commissions is a negatively accelerated decreasing function of categories reaching zero by five or six categories.

2. The maximum number of categories that can be identified in this situation without error is represented by not more than 1 or 2 bits of stimulus information.

3. Neither presentation method nor density affected identification or discrimination except that for the discrimination task, densities greater than one impaired performance.

4. A distinction must be made between information processed as identified categories and as discriminated categories or between perceptual capacity and short term memory capacity. The results indicate that perceptual capacity is the greater of the two.

References

Anderson, N. S., & Fitts, P. M. Amount of information gained during brief exposures of numerals and colors. *J. exp. Psychol.*, 1958, **56**, 362–369.

Christner, C. A., Schutz, H. G., & Ray, H. W. Some factors affecting visual search time for symbols on a large display. Paper presented at the sixty-seventh Annual Convention of the American Psychological Association, Cincinnati, Ohio, 1959.

Green, B. F., & Anderson, L. K. Color coding in a visual search task. *J. exp. Psychol.*, 1956, **51**, 19–24.

Miller, G. A. The magical number seven, plus or minus two: Some limitations on our capacity for processing information. *Psychol. Rev.*, 1956, **63**, 81–97.

Shannon, C. E., & Weaver, W. *The mathematical theory of communication.* Urbana: Univer. Illinois Press, 1949.

Woodworth, R. S., & Schlosberg, H. *Experimental psychology.* New York: Holt, 1954.

(Received April 16, 1960)

Journal of Experimental Psychology
1961, Vol. 61, No. 6, 501–504

MONETARY REWARD AND MOTIVATION IN DISCRIMINATION LEARNING[1]

LOUISE BRIGHTWELL MILLER[2] AND BETSY WORTH ESTES

University of Kentucky

A number of comparative studies have shown that amount of reinforcement is an effective variable in certain classical and instrumental conditioning situations. Experimental attempts to generalize this finding to discrimination tasks have produced conflicting results. Evidence on human *Ss* is particularly limited and also inconsistent (Cantor & Hottel, 1955; Terrell & Kennedy, 1957).

For decisive effects with human *Ss* rather large differences in amount might be necessary because knowledge of results may provide a basic incentive level sufficient to mask small accretions. No-difference results might also be attributable to the possibility that motivation was already at asymptote (Muenzinger, 1934). Moreover, an interaction between motivation and incentive might be expected.

This study represents an attempt to measure the effects of a 50 to 1 differential in monetary reward on human discrimination learning. Motivation was measured by "n Achievement" (n Ach) scores. High n Ach scorers have been shown to be better learners (Lowell, 1952) and better performers (Atkinson & Reitman, 1956; Lowell, 1952).

[1] Based on a study submitted by LBM in partial fulfillment of the requirements for a PhD degree from the University of Kentucky. Sincere appreciation is due James S. Calvin for invaluable advice and criticism in planning the experiment. This research was supported in part by a grant from the University of Kentucky Faculty Research Fund to BWE.

[2] Now at the University of Louisville.

METHOD

Design.—The experiment was a factorial, treatment-by-levels design, with two levels of achievement motivation, and three levels of incentive (knowledge of results, 1¢, and 50¢). The general procedure consisted of obtaining n Ach scores and dividing them at the median. Equal numbers of high and low scorers, drawn at random from the original group, then received the different reward levels during discrimination learning. All three incentive groups were informed of success by a red light. The KOR group (control group) received no additional reward. The 1¢ group and the 50¢ group received the appropriate amounts for each correct choice, and forfeited the same amounts for each error. There were 24 *Ss* in each incentive group, a total of 72.

A preliminary experiment was performed to determine whether the forfeiture of money for errors would constitute a different reward situation from that of reward with no corresponding penalty. Analysis of variance indicated no difference between the two methods.

Since Meyer (1951) has shown that range of rewards is an effective variable, the different incentive levels were offered at separate schools. Because of this deliberate confounding of schools with incentive, matching with respect to socio-economic status was necessary, particularly for the main comparison between the 1¢ and 50¢ groups. Schools were selected on the basis of a recent study by Sparks (Sparks & Ray, 1957). Mean occupational status of fathers of experimental *Ss*, using the same index (Warner, Meeker, & Eells, 1949) was exactly the same in the 1¢ and 50¢ groups.

Subjects.—The *Ss* were males only, taken from the third and fourth grades of the public schools. All *Ss* were between CA 9–0 and 9–11.

n Ach scores.—The method of obtaining n Ach scores was patterned after an experiment by Winterbottom (1953). Eight written stories were obtained from a group of 35–40 boys at each school. Printed themes were presented consecutively. For example, the first suggestion read, "Tell a story about

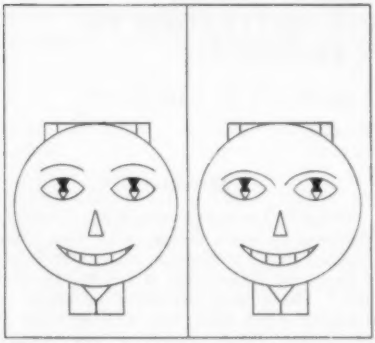

FIG. 1. Stimulus card with high eyebrows on the left.

a mother and her son. They look worried." All eight stories were written in one session, with 4 min. allowed for each story.

The scoring method was modified from that described by McClelland, Atkinson, Clark, and Lowell (1953, Ch. 2, 24). Only two of the suggested categories were used: Achievement-Imagery and Achievement-Thema. Each of these, if present, contributed one point to each story score.

Discrimination task.—The *Ss* were asked to discriminate between two line drawings of faces, which were presented side by side on a card (Fig. 1). The two faces were identical in all respects except the height and spacing of the eyebrows. Two stimulus cards were used, presenting reverse positions of the two faces. Duplicates were used as needed.

Apparatus.—The apparatus was designed for tachistoscopic presentation of stimuli and automatic delivery of reward. On *S*'s side were (from top to bottom): a red signal lamp, viewing window, money slot, tin cup for coins, and selection switch. On *E*'s side a hinged door in the top permitted insertion of cards into the tachistoscope. Below the door were placed: slot for insertion of coins, switch for presetting correct choice, starter switch, red signal lamp, and timer. An incorrect choice merely stopped the timer and removed the card from view. Correct selection lighted red signal lamps on both sides and in the money groups activated a solenoid which released the coin.

Instructions and procedure.—Each *S* was tested individually. The task was presented as one of learning to tell twins apart, Bill from his brother, and to choose Bill every time. Practice was given with a sample pair of square faces, in which the nose of one was larger and higher. Almost all *Ss* recognized this exaggerated difference immediately. Those who did not were shown the difference. The cards were presented in random order with each face appearing equally often on the left and right. The face selected by each *S* on the first trial was rewarded, and was considered correct thereafter. In the money groups, *S* was loaned 10 coins at the beginning, and was told that after the session he would be allowed to keep all that he had received, minus these 10.

Choice was forced after 2 sec. The non-correction procedure was used. Criterion was 16 consecutive correct choices. Trials were massed and sessions terminated after 100. In addition to errors, a record was kept of the time interval between presentation of the stimulus and selection by *S*. Both errors and time were recorded manually.

RESULTS

Overall results are presented in Table 1. Incentive was a significant source of variance, as shown in Table 2. There was no difference between the 50¢ and 1¢ groups, but both these groups made *more* errors than did the KOR group. A two-tailed *t* test showed the difference between the means of the two money groups combined and the KOR group to be significant at better than the .01 level. Of 37 *Ss* who reached criterion performance, 15 were in the KOR group, 10 in the 1¢ group, and 12 in the 50¢ group. These differences, however, are not statistically significant.

n Ach was also a significant source

TABLE 1

MEAN ERROR AND TIME SCORES OF SIX TREATMENT GROUPS

Incentive	Errors			Time (Sec.)		
	High n Ach	Low n Ach	Total	High n Ach	Low n Ach	Total
Knowledge	17.5	24.3	20.9	2.56	2.54	2.55
1 cent	27.6	37.4	32.5	2.10	2.09	2.10
50 cents	27.7	40.3	34.0	2.58	2.32	2.45
Total	24.3	34.0		2.42	2.32	

TABLE 2

ANALYSIS OF VARIANCE OF ERROR SCORES: KOR, 1-CENT, AND 50-CENT GROUPS

Source	*df*	*MS*	*F*
Between groups			
n Achievement (*N*)	1	1,691.68	5.83*
Incentive (*I*)	2	1,246.89	4.30*
$N \times I$	2	102.13	
Within groups	66	290.00	

Note.—Square root transformation of scores produces a slightly smaller *F* for n Ach variance ($P < .05$), but does not alter the probability in regard to incentive variance.
* $P < .025$.

of variance. More errors were made by low n Ach *Ss*, as shown in Fig. 2. The difference between these two means is significant at less than the .025 level, as indicated by a two-tailed *t* test. The number of *Ss* who learned to criterion was significantly larger in the high n Ach group. Of the total, 26 were high n Ach *Ss* and 11 were low. Chi square shows the probability of this difference to be between the .025 and .05 levels.

Analysis of variance of time scores did not reveal any significant variation, and it can be seen from Table 1 that time scores of the six treatment groups were approximately the same.

DISCUSSION

The no-difference results repeatedly found in comparative studies of incentive variation in visual discrimination are further substantiated by these results on human *Ss*. If reward was effective at all, it would be difficult to account for the lack of superiority in the 50¢ group, unless there was some systematic difference which produced better learners in the 1¢ group. But the two money groups did not differ in socioeconomic status; median n Ach scores were the same; mean IQs based on available group scores differed by only 2 points. Correlation between IQ and number of correct trials proved to be low, $r = .19$.

The clear inferiority of the reward groups was an unexpected result, unaccountable for by theory or previous empirical evidence. The implication that monetary incentive had a detrimental effect cannot be substantiated from this study alone, although observations made during the experiment suggested that preoccupation with the money (counting, gloating, worrying) constituted a sort of interpolated task.

On the other hand, any incentive effect—whether facilitating or detrimental—might be expected to interact with motivation. Yet significantly poorer learning with money was the case for both high and low n Ach *Ss*. The statistical analysis revealed no significant interaction between n Ach and incentive (Table 2), although it appears from inspection that the two poorest groups were low n Ach *Ss* who received money.

Motivation effects are apparent, in that low n Ach *Ss* made significantly more errors than high n Ach *Ss*. The relationship between n Ach and IQ does not appear important, $r = .24$. This result extends the generality of the n Ach score as a measure of the need to achieve on a discrimination task.

It is interesting that of those *Ss* who learned, less than half were able to verbalize the relevant cue. Among the

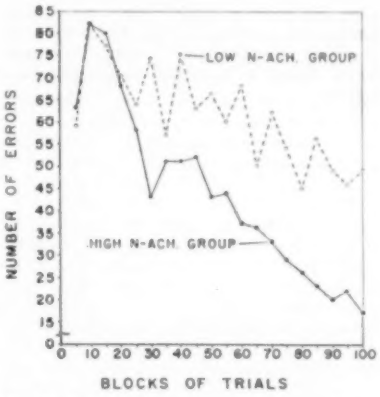

FIG. 2. Effect of n Achievement on acquisition of a visual discrimination habit.

statements made by these *Ss* were: "Bill's mouth is straighter"; "Bill's nose is littler"; "Bill had the most teeth." None of these differences actually existed, of course, yet these *Ss* had made at least 16 successive correct choices, the chance probability of which is so small as to be negligible. They had learned to discriminate without identifying the locus of difference. There were no significant differences in errors or trials for *Ss* who could identify the relevant cue, as compared with those who could not.

Summary

An experiment was performed to investigate the effects of different levels of monetary incentive and n Achievement on the acquisition of a visual discrimination habit. Groups of 9-yr. boys learned to discriminate between drawings of faces which differed only in the height and spacing of the eyebrows. Pictures were presented tachistoscopically for 2-sec. intervals. The *Ss* received after each correct choice: 50¢, 1¢, or only knowledge of results.

The major results were: (*a*) There was no difference in number of errors made by *Ss* rewarded with 50¢ pieces and those given pennies. (*b*) Groups who received money made significantly *more* errors than those receiving only knowledge. (*c*) Low n Ach *Ss* made significantly more errors than high n Ach *Ss*. (*d*) There were no significant differences among experimental groups in time required to respond. These results were interpreted as substantiating the no-difference results repeatedly found in comparative studies of incentive variation on discrimination learning.

References

Atkinson, J. W., & Reitman, W. R. Performance as a function of motive strength and expectancy of goal-attainment. *J. abnorm. soc. Psychol.*, 1956, **53**, 361–366.

Cantor, G. N., & Hottel, J. V. Discrimination learning in mental defectives as a function of magnitude of food reward and intelligence level. *Amer. J. ment. Defic.*, 1955, **60**, 380–384.

Lowell, E. L. The effect of need for achievement on learning and speed of performance. *J. Psychol.*, 1952, **33**, 31–39.

McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. *The achievement motive*. New York: Appleton-Century-Crofts, 1953.

Muenzinger, K. F. Motivation in learning: I. Electric shock for correct response in the visual discrimination habit. *J. comp. Psychol.*, 1934, **17**, 267–277.

Myer, D. R. The effect of differential rewards on discrimination reversal learning by monkeys. *J. exp. Psychol.*, 1951, **41**, 268–274.

Sparks, P. E., & Ray, L. C. An evaluation of two methods of teaching reading. *Elem. sch. J.*, 1957, **57**, 386–390.

Terrell, G., Jr., & Kennedy, W. A. Discrimination learning and transposition in children as a function of the nature of the reward. *J. exp. Psychol.*, 1957, **53**, 257–260.

Warner, W. L., Meeker, M., & Eells, K. *Social class in america: A manual of procedure for the measurement of social status*. Chicago: Science Research Associates, 1949.

Winterbottom, M. The relation of childhood training in independence to achievement motivation. Unpublished doctoral dissertation, University of Michigan, 1953.

(Received April 25, 1960)

Journal of Experimental Psychology
1961, Vol. 61, No. 6, 505–507

THE EFFECT OF SEPARATION BETWEEN TEST AND INDUCING FIELDS ON BRIGHTNESS CONSTANCY[1]

BRUCE DUNN AND H. LEIBOWITZ

University of Wisconsin

Historically (Hering, 1920; Woodworth & Schlosberg, 1954), discussions of the mechanisms underlying brightness constancy have considered the possible role of simultaneous brightness contrast. Such analyses have been hindered, however, by the lack of an unequivocable specification of the conditions under which contrast occurs experimentally (Graham, 1934). More recently, a number of studies on contrast have appeared which provide a more precise specification of contrast and which permit consideration of the contribution of contrast to brightness constancy. Diamond (1953, 1955) and others (Heinemann, 1955; Leibowitz, Mote, & Thurlow, 1953) determined that the contrast effect between two simultaneously visible fields is primarily unidirectional. The brighter field (referred to as the inducing field) inhibits the dimmer (referred to as the test field), while the dimmer field has little if any effect on the brighter. Utilizing this specification, it has been determined that although some constancy is present in the absence of contrast as here defined, the presence of contrast markedly increases the tendency toward constancy. Such results support the assumption that contrast is indeed a major variable contributing to the phenomenon of brightness constancy. The feasibility of this assumption was strengthened by two subsequent experiments in which reduction of exposure duration increased both contrast (Chinetti, 1957) and constancy effects (Leibowitz & Chinetti, 1957; Leibowitz, Chinetti, & Sidowski, 1956). The purpose of the present study is to provide a further test of the relationship between contrast and constancy by manipulation of the separation between test and inducing fields, a procedure which has previously been shown to result in a decrease in simultaneous brightness contrast (Leibowitz, Mote, & Thurlow, 1953).

[1] Supported in part by Grant M1090 from the National Institute of Mental Health of the National Institutes of Health, United States Public Health Service. Computations were aided by facilities supplied by the University of Wisconsin Numerical Analysis Laboratory with funds provided by the Wisconsin Alumni Research Foundation.

APPARATUS AND PROCEDURE

The apparatus has previously been described in detail (Leibowitz, Myers, & Chinetti, 1955). The *S* is seated in front of the partition dividing two adjacent, equal sized chambers. In the left chamber a square "grey" test field of 36.2% reflectance (No. 12 Munsell Grey) subtending a visual angle of 1° at a viewing distance of 6 ft. is viewed under various levels of illuminance produced by sources not visible to *S*. In the right chamber, which is otherwise dark, *S* views a photometric field of the same dimensions and located in a position comparable to that of the test field. The *S*'s task is to adjust the luminance of the photometric field so that it matches the test field. An inducing field is provided by an 86% reflectance "white" subtending 6° on each side and placed just above and contiguous to the test field (designated as 0° separation), and at separations of 6° and 12° between the nearest edges of the fields. The *S* observed binocularly at all times with unrestricted head and eye movements. The walls and floors of the chambers were painted "flat" black (reflectance 1.78%).

Data were obtained from 16 *S*s, volunteers from elementary psychology classes, under

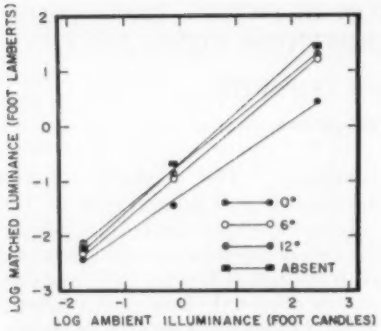

FIG. 1. Log matched luminance of the comparison field as a function of the log illuminance incident on the test stimuli with separation between test and inducing fields as parameter. (On this plot, a slope of unity indicates agreement with the "additivity law of luminance." A slope of zero represents the "law of brightness constancy.")

three levels of illuminance (−1.75, −0.16, and 2.44 log ft-c) in the absence of an inducing field, and with the inducing field present at three separations. Order of presentation of illuminance level and inducing field conditions was determined by counterbalanced squares.

RESULTS

The mean data for all *S*s are presented in Table 1 and plotted as the log of the mean matched luminance as a function of the log ambient illuminance in Fig. 1. On this plot, a horizontal line represents a prediction in terms of brightness constancy which assumes that the same match would be obtained between the test and photometric fields at all levels of illuminance. At the other theoretical extreme the additivity law of luminance, under which perceived brightness is determined solely by the quantity of incident luminous energy, is here represented by a straight line of unit slope. As is typical in such studies, the matches lie between the two theoretical extremes. However, the extent to which the data approach the theoretical brightness constancy function is related to the separation between the test and inducing field. For adjacent test and inducing fields both contrast, as indicated by a shift of the entire function to lower matched luminance values, and constancy, as indicated by a lesser slope (0.67) are greatest. With increasing separation and in the absence of the inducing field, both contrast and constancy decrease. The slopes of the functions for 6° and 12° of separation and in the absence of the inducing field are 0.84, 0.85, and 0.88, respectively. The most marked effect of separation occurs when comparing the adjacent inducing field with 6° of separation. Separation beyond 6° as in the previous contrast study, has only a negligible effect. Analysis of variance indicated that both luminance and separation are significant at beyond the .001 level.

DISCUSSION

The results of the present experiment are in agreement with the assumption that simultaneous contrast can play a significant role in brightness constancy. This reasoning is based on the assumption that when two behavioral measures covary, their underlying mechanisms share common elements. It has previously been demonstrated that both contrast and constancy increase with reduc-

TABLE 1

MEAN LOG MATCHED LUMINANCE IN FOOT-LAMBERTS AS A FUNCTION OF ILLUMINANCE LEVEL WITH SEPARATION BETWEEN TEST AND INDUCING FIELD AS THE PARAMETER

Separation in Degrees of Arc	Log Ambient Illuminance (Foot-Candles)		
	−1.75	−0.16	2.44
0	−2.42	−1.45	0.44
6	−2.32	−0.99	1.24
12	−2.19	−0.92	1.26
Inducing field absent	−2.25	−0.78	1.41

tion of exposure duration, and with an increase of the luminance of the inducing field relative to that of the test field. The present study represents another such measure of covariance, i.e., a decrease in both measures with separation between test and inducing fields.

Such results, however, imply only that simultaneous contrast is *one* of the variables contributing to constancy as some tendency toward constancy is observed in the absence of contrast, e.g., the condition in the present study in which the inducing field was absent. There are a number of possibilities. Changes in pupil diameter could be expected to decrease the difference in perceived brightness of objects viewed under different illuminance levels. This would tend to decrease the slopes of the matched luminance-ambient illuminance functions, and to induce some tendency toward constancy. With familiar objects, especially outside of the laboratory, memory and context probably play an important role (Bolles, Hulicka, & Hanly, 1959; Hering, 1920; Woodworth & Schlosberg, 1954). The particular significance of contrast is due to the fact that it can be readily demonstrated in lower forms both behaviorally (von Buddenbrock, 1952) and electrophysiologically (Ratliff, Miller, & Hartline, 1958). Thus, contrast would appear, by virtue of its biological generality, to be a basic mechanism subserving the brightness constancy phenomenon.

Summary

Luminance matches were obtained between a 1° square "grey" test field viewed under various levels of illuminance and a photometric field of the same size. Data were obtained with the test field viewed against a "black" background, and with a "white" inducing field introduced at various degrees of separation from the test field.

The tendency toward brightness constancy was observed under all experimental conditions, but increased markedly as the separation between test and inducing fields was decreased.

The results are interpreted as supporting the point of view that simultaneous brightness contrast is one of the important mechanisms underlying the phenomenon of brightness constancy.

References

Bolles, R. C., Hulicka, I. M., & Hanly, B. Colour judgment as a function of stimulus conditions and memory color. *Canad. J. Psychol.*, 1959, **13**, 175–185.

Chinetti, P. J. The effect of reduced exposure duration on simultaneous brightness contrast. Unpublished doctoral dissertation, University of Wisconsin, 1957.

Diamond, A. L. Foveal simultaneous brightness contrast as a function of inducing and test-field luminances. *J. exp. Psychol.*, 1953, **45**, 304–314.

Diamond, A. L. Foveal simultaneous contrast as a function of inducing-field area. *J. exp. Psychol.*, 1955, **50**, 144–152.

Graham, C. H. Vision: III. Some neural correlations. In C. Murchison (Ed.), *Handbook of general experimental psychology*. Worcester, Mass.: Clark Univer. Press, 1934.

Heinemann, E. G. Simultaneous brightness induction as a function of inducing- and test-field luminances. *J. exp. Psychol.*, 1955, **50**, 89–96.

Hering, E. *Grundzüge der Lehre vom Lichtsinn*. Berlin: Springer, 1920.

Leibowitz, H., & Chinetti, P. Effect of reduced exposure duration on brightness constancy. *J. exp. Psychol.*, 1957, **54**, 49–53.

Leibowitz, H., Chinetti, P., & Sidowski, J. Exposure duration as a variable in perceptual constancy. *Science*, 1956, **123**, 668–669.

Leibowitz, H., Mote, F. A., & Thurlow, W. R. Simultaneous contrast as a function of separation between test and inducing fields. *J. exp. Psychol.*, 1953, **46**, 453–456.

Leibowitz, H., Myers, N. A., & Chinetti, P. The role of simultaneous contrast in brightness constancy. *J. exp. Psychol.*, 1955, **50**, 15–18.

Ratliff, F., Miller, W. H., & Hartline, H. K. Neural interaction in the eye and the integration of receptor activity. *Ann. NY Acad. Sci.*, 1958, **74**, 210–222.

von Buddenbrock, W. *Vergleichende Physiologie*. Band I. *Sinnesphysiologie*. Basel: Birkhäuser, 1952.

Woodworth, R. S., & Schlosberg, H. *Experimental psychology*. (Rev. ed.) New York: Holt, 1954.

(Received May 2, 1960)

Journal of Experimental Psychology
1961, Vol. 61, No. 6, 508–509

SUPPLEMENTARY REPORT: SPATIAL GENERALIZATION OF VOLUNTARY RESPONSES UNDER TWO TECHNIQUES OF STUDY AND TWO LEVELS OF ANXIETY

BRYAN D. DIXON[1] AND DELOS D. WICKENS

Ohio State University

Bilodeau, Brown, and Meryman (1956) reported a method of studying stimulus generalization wherein *Ss* were requested to react manually as rapidly as possible to the central one of a group of seven equidistant lights but not to the others. Brown, Clark, and Stein (1958) modified this from a reaction time situation to one where the illumination of a light signified that this particular "horse" had won a race and *Ss* were permitted to bet "win" or "lose" on each horse on individual trials. The central light flashed on 80% of the trials, the others, 20%. These will be referred to as Exp. RT and Exp. HR, respectively. Both experiments report a double-winged gradient of reactions or win bets, decreasing from the center to the periphery. In the present experiment the above mentioned methods were employed while the *Ss* were administered shock at various times during the experiment.

Method.—There were 25 *Ss* (college students) in each of four groups, Group S (shock) and Group NS (no shock) for each situation. The apparatus and the general procedure attempted to duplicate the above-mentioned experiments, except that 2.0-ma. shocks were given at random times to the shock groups. The shocks were not associated with any particular response.

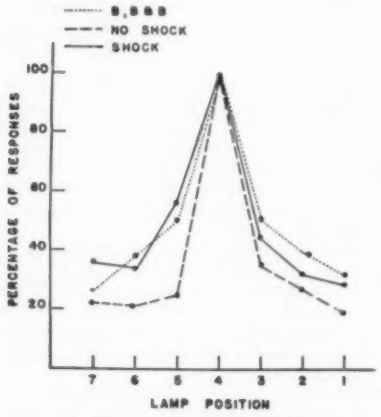

FIG. 1. Results for Exp. RT.

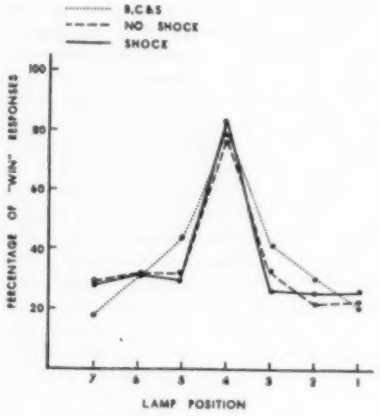

FIG. 2. Results for Exp. HR.

Results and discussion.—The results are presented graphically in Fig. 1 and 2. These figures also show the data of the studies being replicated. In Exp. RT, Group S made significantly more false responses to the peripheral lamps than did Groups NS ($P < .01$). Under both conditions there is some evidence of a generalization gradient. In Group S the light adjacent to the center differed from the extreme peripheral light on both wings, but it did not do so on the left wing for Group NS ($P < .05$ was considered evidence of a difference between points). In Exp. HR no significant difference was found between the overall frequency of "win" responses for Groups S and NS. In Group S no peripheral point differed significantly from any other point, producing no evidence for a generalization gradient. Some evidence for a gradient is found in Group NS, for the response frequency at Lamp 3 is significantly higher than at Lamps 1 and 2 ($P < .02$).

The experiment was designed to test the

[1] Now with the Quartermaster Corps, United States Army.

hypothesis that anxiety generated by shock would increase the predominant response tendency through its multiplicative relationship to habit strength (Spence, 1956). Experiment RT offers support for this hypothesis for false responses are more frequent under the shock than the no shock condition.

Although at first blush the results of Exp. HR may seem to contradict this hypothesis, such is not the case. Indeed the experiment was begun with the prediction that anxiety should, according to this theory, *decrease* the number of "win" responses. It should do so because the response "lose" could be reinforced 80% of the time if given to the peripheral light, and, hence this response should be the dominant response for the peripheral lights. Since there is no significant difference between Groups S and NS the theory is not supported in this instance. It receives some support from the fact that introduction of shock completely eliminated any gradient, whereas there is some evidence of a gradient on the left wing for Group NS. Parenthetically, it is rather surprising that the Horse Race should produce a gradient, for *S* can win by saying "lose" to the peripheral lights. Since there are 6 peripheral lights where losing predominates in frequency, and only the central light where winning predominates, it might be expected that generalization would take the form of suppressing the frequency of win responses for the central light, but in both groups the guessing frequency is close to the objective probability, 83% and 77% for Groups S and NS, respectively. It is also interesting to note that "win" responses are given at greater than the objective frequency for the peripheral lamps. Apparently, *Ss* would rather win by betting on a winner than a loser.

References

Bilodeau, E. A., Brown, J. S., & Meryman, J. J. The summation of generalized reactive tendencies. *J. exp. Psychol.*, 1956, 51, 293–298.

Brown, J. S., Clarke, F. R., & Stein, L. A new technique for studying spatial generalization with voluntary responses. *J. exp. Psychol.*, 1958, 55, 359–362.

Spence, K. S. *Behavior theory and conditioning.* New Haven: Yale Univer. Press, 1956.

(Received December 27, 1960)

Journal of Experimental Psychology
1961, Vol. 61, No. 6, 509–510

SUPPLEMENTARY REPORT: DIRECTIONAL EFFECTS AND SEX IN KINESTHETIC AFTEREFFECTS[1]

PAUL BAKAN, RICHARD THOMPSON, AND GAIL WILDES

Michigan State University

The kinesthetic aftereffect (KAE) is a change in apparent width of a standard stimulus after exposure to an inspection (I) stimulus, wider or narrower than standard. Though KAEs in opposite direction are produced by varying direction of difference between standard and I-stimulus, little is known about the effect of direction of difference on size of KAE. Köhler and Dinnerstein (1947) report that exposure to a 2.5 in. I-stimulus decreases a narrower test width more readily than it enlarges a wider one. This finding is based on very few *Ss* and is not statistically evaluated.

This study compares the effect of I-stimuli, wider and narrower than standard, with absolute difference between I-stimulus and standard held constant. The design also allows for analysis of sex differences in KAE, which have been reported (Barthol, 1958; Spitz & Lipman, 1960; Wertheimer, 1954). Male and female *Ss* and a male and female *E* were used.

Method.—A group of 172 volunteer undergraduate *Ss* (M = 18–28 yr.; *F* = 17–23 yr.) were assigned to eight groups in a 2 × 2 × 2 design with two I-conditions, wider (2.5 in.) and narrower (0.5 in.) than standard (1.5 in.), two *S* sexes, and two *E* sexes. Four groups of 25 *Ss* were tested by the male *E* and four groups of 18 *Ss* were tested by the female *E*.

Apparatus consisted of a standard stimulus, two I-stimuli, and a variable width wedge-shaped scale ranging from $2\frac{1}{4}$ in. to 1 in. and changing at the rate of $\frac{1}{32}$ in. per inch. These were made of wood and were mounted on tables at a height of 36 in. The *S* was blindfolded and led between two tables. The standard and the I-stimulus were on the side of *S*'s nonpreferred hand, with the scale on the side of the preferred hand. Stimuli were held between the thumb and middle finger.

[1] This investigation was supported by a research grant, G5822, from the National Science Foundation.

TABLE 1

MEAN KAE BASED ON DIFFERENCE BETWEEN PRE-I AND POST-I JUDGMENTS OF STANDARD (INCHES)

Ss and *Es*	Inspection Stimulus					
	Wide			Narrow		
	Pre-I	Post-I	Diff.	Pre-I	Post-I	Diff.
Male *Ss*						
Male *E*	1.52	1.44	−.08	1.52	1.64	.12
Female *E*	1.45	1.36	−.09	1.53	1.64	.11
Female *Ss*						
Male *E*	1.50	1.41	−.09	1.49	1.64	.15
Female *E*	1.53	1.47	−.06	1.57	1.70	.13
All	1.50	1.42	−.08	1.53	1.66	.13

Following four practice judgments *S* made four pre-inspection judgments which involved finding a place on the scale, apparently equal in width to the standard. Judgments were made in an ADDA order from points .25 in. narrower and wider than objective equality. The *S* then moved his nonpreferred hand back and forth along the I-stimulus for 90 sec., while the other hand was held palm down with fingers together on a table. After the I-period *S* again made four judgments of the width of the standard. The difference between the means of the pre-inspection and postinspection judgments constitutes the measure of KAE.

Results and discussion.—The mean KAE for each group is presented in Table 1 together with the measures from which it is derived. Results of an analysis of variance of the data indicate that the difference between the wide and narrow I-stimulus is significant at the .01 level ($F = 6.91$, $df = 1/164$, error mean square = 212). The narrow I-stimulus produces a greater change in apparent width of the standard than does the wide I-stimulus. These results agree with and extend the results of Köhler and Dinnerstein in showing that absolute difference between the I-stimulus and standard does not in itself determine the size of the KAE. It should be noted that despite identity of absolute differences between the I-stimuli and standard, their ratios are not identical for the two I-conditions. The effect of direction of difference between the I-stimulus and the standard may be related to the phenomenal difference between the stimulus differences. The psychophysical relationship (Tomlinson, 1960) suggests that the difference between stimuli of 0.5 in. and 1.5 in. may be phenomenally greater than that between 1.5 in. and 2.5 in. If there is a relationship between size of KAE and apparent difference between the I-stimulus and the standard, such relationship must hold only within limits, since it has been shown that the KAE is subject to the distance paradox (Charles & Duncan, 1959; Krauskopf & Engen, 1960) which produces a diminution of KAE when the difference between the I-stimulus and the standard is sufficiently large.

Subject sex differences were not significant; nor were *E* sex differences or any of the interactions significant. Sex differences in KAE have been reported by others, though not universally (Charles & Duncan, 1959; McEwen & Rodger, 1960).

REFERENCES

BARTHOL, R. P. Individual and sex differences in cortical conductivity. *J. Pers.*, 1958, **26**, 365–378.

CHARLES, J. P., & DUNCAN, C. P. The distance gradient in kinesthetic figural aftereffect. *J. exp. Psychol.*, 1959, **57**, 164–170.

KÖHLER, W., & DINNERSTEIN, D. Figural aftereffects in kinesthesis. In, *Miscellanea psychologia Albert Michotte*. Paris: Libraire Philosophique, 1947. Pp. 196–220.

KRAUSKOPF, V., & ENGEN, T. The distance paradox in kinesthetic aftereffects. *Amer. J. Psychol.*, 1960, **73**, 298–301.

MCEWEN, P., & RODGER, R. S. Some individual differences in figural aftereffects. *Brit. J. Psychol.*, 1960, **51**, 1–8.

SPITZ, H. H., & LIPMAN, R. S. Reliability and intercorrelation of individual differences on visual and kinesthetic figural after-effects. *Percept. mot. Skills*, 1960, **10**, 159–166.

TOMLINSON, J. T. Weber's law and thickness determined tactually. *Amer. J. Psychol.*, 1960, **73**, 316–317.

WERTHEIMER, M. Constant errors in the measurement of figural after-effects. *Amer. J. Psychol.*, 1954, **67**, 543–546.

(Received November 1, 1960)